

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIYSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVWCGVSALSMLTCSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFLTYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
 GTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTC
 CTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTG
 GATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAG
 GTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATCTTTGAAATCTTAGG
 AGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGAATCTGCTGATCCTGG
 TTTTCATGGTGCTTTTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCATAAACAACGA
 CTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCC
 CATTCTCAGCCCAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAG
 TGACTCTCATGGCTCTTCTTTCTGGATTGGTGCTGTCAACTGCCCATACACTTACATGTCTTACTTC
 CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
 CATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCAGAAGGGGGAGTGCATAACA
 AACCATCAGGTTTCTGGGAATGATAAAAAGTGTACCACTTCAGCATCAGGAAGTAAAAATCTTACT
 CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGGAACAGCTGATCT
 ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTCAAGGGGAAATATTTAATTTTCTTGTT
 ACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTT
 GGGAAAACGGATCCTGTCAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGT
 GAAGTTTGGTCCCAACACATTTCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGAT
 TGCTGATCACTCTTACCAAGTCTTTTATGCCATCTCTAGCAGTAAGTCTCCAATGTCATTGTCCTG
 CTATTAGCACAGATAATGGGCATGTACTTTGTCTCTCTGTGCTGCTGATCCGAATGAGTATGCCTTT
 AGAATACCGCACCATATAATCACTGAAGTCTTTGGAGAAGTGCAGTTCAACTTCTATCACCGTTGGTTTG
 ATGTGATCTTCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCA
 GAGAAGCAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAATTTA
 GATATAAGAGGGGGGAAAAATGGAACAGGGCCTGACATTTATAAACAACAAAATGCTATGGTAGC
 ATTTTTACCTTCATAGCATACTCCTTCCCCGTGAGGTGATACTATGACCATGAGTAGCATCAGCCAG
 AACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG
 TGTAGAGGCGGAGAGGAGCCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT
 CTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTAAAGGTTACATGGAAAAGGTTATAGCTTTG
 CCTTGAGATTGACTCATTAAAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAGGGCGGCCGCG
 ACTCTAGAGTCGACCTGCAGAAGCTTGCCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGBP
FPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDALAEELSRO
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLLAQIMGMV
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSAISSILFLYLAHKQAEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTC'CGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
 GTTTCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTCCATCCAGGT
 GTCATGCAGAATTATGGGGATCACCC'TGTGAGCAAAAAGGCGAACCAGCAGCTGAATTTACAG
 AAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTTGAAACAGCC
 TTGAAAGCTAGCTTTGAACTTGCAGCTATGGCTGGGTTGGAGATGGATTTCGTGGTCATCTCTAG
 GATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCTGATTTGGAAGGTTCCAGTGA
 GCCGACAGTTTGCAGCCTATGT'TACAAC'TCATCTGATACTTGGACTAACTCGTGCATTCCAGAA
 ATTATCACCACCAAAGATCCCATATTTCAACTCAAACGCAACAAACAAGAAATTTATTGT
 CAGTGACAGTACCTACTCGGTGGCATCCCC'TACTCTACAATACCTGCCCTACTACTACTCCTC
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTGATTTGTGTACAGAAGTTT'TATG
 GAACTAGCACCATGTCTACAGAACTGAACCATTTGTGAAAATAAAGCAGCATTCAGAATGA
 AGCTGCTGGGTTTGGAGGTGTCCCCACGGCTCTGCTAGTGCTTGCTCTCCTCTTCTTGGTGTG
 CAGCTGGTCTTGGATTTTGTCTATGTCAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAA
 CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAA
 TGAGGAATCAAAGAAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGC
 GATGCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTTT
 CATGCTCCTTACCCTGCCCCAGCTGGGGAATCAAAGGGCCAAAGAACCAAAGAAGAAAGTCCA
 CCCTTGGTTCCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAAT
 GCCCTTCTCCTTATTGTAACCCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCT
 TTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTT'TGCAAAGTGAAGGAC
 CTAAACATCTCATCAGTATCCAGTGGTAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTTG
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGCAGCTCAGACCC'TTCTTCA
 GCTCTGAAAGAGAAACACGTATCCCACTGACATGTCTTCTGAGCCCGGTAAGAGCAAAAGAAT
 GGCAGAAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA
 GCTAAATAAAGAAATAGAACAAAGGCTGAGGATACGACAGTACACTGTGAGCAGGGACTGTAAAC
 ACAGACAGGGTCAAAGTGT'TTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACA
 CTTACTTTTCTGGTCTCTACCACTGCTGATATTTTCTCTAGGAAATATACTTTTACAAGTAACA
 AAAATAAAACTCTTATAAATTTCTATTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCACAAACATTTGCTGAATAGCTACTATATGTC
 AAGTGCTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATGCACATAGTAG
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTGTATTTCTAGCTTATCTACTTCCAACTAAT
 TTTTATTTTGTGCTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAAT
 TATTATTAAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAAGCACATTTTAAAGTGCC
 ATTAACAAATGTATCACTAGCCCTCCTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATT
 TGTGACAAAAAATTAAAGCATTTAGAAAACCTT

FIGURE 6

MARCFSLVLLLSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAAFNEAAGFGGVP TALLVLALLFFGAAAGLGFCYVK
RYVKAFFPTNKNQOKEMIETKVVKEEKANDSNPNEESKKT DKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCCCCGCCACCGCGCGCTCCCGCATCTGCACCCGCGAGCCCGGC
GGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCCGAGCGCAACTCGGTCCAGTCGGGGCGG
CGGCTGCGGGCGCAGAGCGGAGATGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGCTGGCGG
CGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCG
GCTCTCAGCTACCCGCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAAGTAT
GGAGGACACGAGCACAATGCGCAGCGCGGTGGAAGAGATGGAGGCGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTACCTCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACAGAC
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGCTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACAAAATGGCCACCAGGGGCGAGCAATGGGACCATTCTGTGACA
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCCGTGCTGTTCCCTGTGTGC
ACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTA
CATCTTCTTCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTCAGCT
CCCCAGGCTGTTCTCCAGGCTTACAGTCTGGTGTCTTGGGAGAGTCAGGCAGGGTTAACTGCA
GGAGCAGTTTGGCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAACAATGTGGAGTCTCCCTC
TGATTGGTTTTGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAACATCAACCTGGCAAAAATG
CAACAAATGAATTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGC
AGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCT
CTCCACTACCCACACCAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAATAATCTCACATCCCTCTAAAAGTAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATG
TTTTCAGGTGTCATGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTTAGAA
ATCAAGCATAAATCACTTCACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

[illegible]

Signal sequence:
amino acids 1-19

amino acids 96-100, 106-110, 121-125, 204-208

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

amino acids 202-208, 217-223

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
GGCCACCTTGTGAACCTCTCGTGCCCGAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTCTTCTGGACCTT
AACTGGGTACTGGCCCTGGGCCAATGCGTCTCGCTGGAGCCTTTGCCTCCTTCTACTGGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
ACACTGGGTCATTGGCATTGGAGCCCTCATCCTGACCCTTGTGCAGATAGCCCGGGTCATCTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGAGAACCCTGTAGCCCGCTGCATCATGTGCTGTTT
CAAGTGCTGCCTCTGGTGTCTGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC
ATTGTCAGGGTGGTCGCTCGTGGACAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
GGTCGGAGGCGTGGGGTCTGTCTTCTTTTTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGGCTAT
GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTTCTCTGCTTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA
AGATTCTGGGCAAGAAGAACGAGGCGCCCCGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCCAACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT
CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCGGTGGCTCACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAAACATGGTG
AAACCTCCGTCTCTATTAATAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCTATCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTCAGTGAGCCGA
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAAACAAA
AAGATTTTATTAAAGATATTTGTGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLEFWTL
 NWVLALGQCVLGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
 EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
 IVRVVLDKVTDLLLFFGKLLVVGGVVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
 VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCGCCGCGCCCGGCGCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCCTGC
 TCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCCG
 CAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCTCTTCTGGGGTGCTGGTGTTCCA
 TCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGGCC
 GGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGCTTACCG
 CATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTTTACCCTGCTCATGCTCTGCGTGAGCAGCA
 GCCGGGACCCCGGGCTGCCATCCAGAATGGGTTTTGGTTCTTTAAGTTCTGATCCTGGTGGGCCTC
 ACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTTACCAACATCTGGTTCTACTTCGGCGTCGTGGG
 CTCCTTCTCTTCATCCTCATCCAGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACACGCGGT
 GGCTGGGCAAGGCCGAGGAGTGCGATTCCCGTGCTGGTACGCAGGCCTTCTTCTTCACTCTCCTC
 TTCTACTTGCTGTGATCGCGGCGGTGGCGCTGATGTTTCATGTACTACACTGAGCCAGCGGCTGCCA
 CGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCATCGCTGCTGTCTCTGC
 CCAAGGTCCAGGACGCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTATCACCTCTACACCATG
 TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCATTTGCCAACCCAGCT
 GGGCAACGAGACAGTTGTGGCAGGCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCGAGCATTG
 TGGGCCTCATCATCTTCTCTGTGCACCTCTTCATCAGTCTGCGCTCCTCAGACCACCGGCAGGTG
 AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGTT
 GGCAGCCTGTGAGGGCCGGGCCTTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCC
 ACTTCTGCCTGGTGCTGGCCTCACTGCACGTCATGATGACGCTCACCAACTGGTACAAGCCCGGTGAG
 ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCT
 CCTCTACCTGTGGACCTTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCTCA
 CAGCCTGCCATCTGGTGCTCCTGCCACCTGGTGCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTC
 CCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTGCCCCCTGAGCCGGGC
 CTTCTAGTCGTAGTGCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC
 ACCCACACGGTGGAGCTGCCTCTTCTTCCCCCTCCTCCTGTTGCCATACTCAGCATCTCGGATGAA
 AGGGTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACCTCCACCACAG
 TGGGGCATCCGGCACTGAAGCCCTGGTGTTCCTGGTCACGTCCCCCAGGGGACCTGCCCCCTTCTCTG
 GACTTCGTGCCTTACTGAGTCTCTAAGACTTTTTCTAATAAACAAGCCAGTGCCTGTAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRIFTFFFLGLVLSIIMLSPGVESQL
 YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFTLMLCVSSSRDPRAAIQ
 NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQRWLKAE
 ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVSIAAVLPKV
 QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
 VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVAACEGRAFDNEQDGVITYSY
 SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVVKICASWAGLLLYLWTLVAPLLLRNRD
 FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
 272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
CTCATGGCCAGGATTGAGTCTTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTTGTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCAATATGCTGTGTGCAG
ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG
TGATCCTTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTCATTATC
CTTGCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTTTCAAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG
GTCAGTTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTTAAATGTGAAAAACCCCTCACAGAAAGTC
ATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTATCCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAATAAGGTGAAGAAAGTCTTGCTGTGATTCTTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTATTTCATTCAC
AGAATGGAATTTTTTTGTTTCATGCTCAGATTTATTTTGTATTTCTTTTTTAACACTCTACATT
TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTAAAAAGTGAATAAAATCTG
ACATGTCAATGTGGCTAGTTTTATTTTCTGTTTTGCATTATGTGTATGGCCTGAAGTGTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTT
TTATCATGAAATCATGTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC
ACAAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTGAGCAATTTCAATTAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWIETWFLDFKVLQPQAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGGCCGACCCGCCAGGAAAGACTGAGG
CCGCGGCCTGCCCCGCGCGGCTCCCTGCGCCGCGCGCCCTCCCGGGACAGAAGATGTGCTCCAG
GGTCCCTCTGCTGCTGCGCTGCTCTGCTACTGGCCCTGGGGCTGGGGTGCAGGGCTGCCCAT
CCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCACGGTGCCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC
AGGCAGCTTTGCGGCCTGCGGGCCTGCAGCTCCTGGACCTGTACAGAACCAGATCGCCAGCC
TGCCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG
CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGGCCTCGAGCGCCTCTACCTGGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC
AGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGCCTGCCCCGCTGCTGCTGGACCTCAGC
CACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTGCCAACGTGGAGGCGCTGCGGCT
GGCTGGTCTGGGGTGCAGCAGCTGGACGAGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACC
TGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCTCCGGGGCTGACG
CGCCTGCGGCTGGCGGCAACACCCGCATTGCCAGCTGCGGCCCGAGGACCTGGCCGGCTGGC
TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCGCTGCTGGCGACCTCTCGGGCC
TCTTCCCCCGCTGCGGCTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGC
TGGTTTGGCCCCCTGGGTGCGCGAGAGCCACGTCACTAGCCAGCCCTGAGGAGACGCGCTGCCA
CTTCCCGCCCAAGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAG
CCACCACCACACAGCCACAGTGCCACACGAGGCGCTGGTGGCGGAGCCACAGCCTTGTCT
TCTAGCTTGGCTCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCCCCTC
CACTGCCCCACCGACTGTAGGGCCTGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCA
ATGGGGGCACATGCCACCTGGGGACACGGCACCACTGGCGTGCTTGTGCCCGAAGGCTTACG
GGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACAGTCACGCCGAG
GCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC
AGCGCTACCTCCAGGGGAGCTCCGTGCGAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG
GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCTCGCTGAGTACACGGTCACCCA
GCTGCGGCCCCAACGCCACTTACTCCGTCTGTGTATGCCTTTGGGGCCCGGGCGGGTGCCGAGG
GCGAGGAGGCTGCGGGGAGGCCATACACCCAGCCGTCCACTCCAACCACGCCCCAGTCACC
CAGGCCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCGCCCTGGCCGCGGTGCTCCTGGCCGC
GCTGGCTGCGGTGGGGGACGCTACTGTGTGCGGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGG
ACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCCTGGAAGTGGAGGGAGTGAAGGTCCCCTTGGAG
CCAGGCCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC
ACTCATGGGCTTCCAGGGCCTGGCCTCCAGTCACCCCTCCACGCAAGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGGCCGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCC
ACACCACGTAAGTCTCAGTCCCAACCTCGGGGATGTGTGAGACAGGGCTGTGTGACCACAGCT
GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC
CTAACGTCCCCAGAACCAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC
CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCCCTGGGTCTGCTGGGCTCTCCAC
TCCAGGCGGACCCTGGGGGCCAGTGAAGGAAGTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGC
TTTAGGAACATGTTTTGCTTTTTTAAATATATATTTATAAGAGATCCTTTCCCATTTTATCT
GGGAAGATGTTTTTCAAACTCAGAGACAAGGACTTTGGTTTTTTGTAAGACAACGATGATATGAA
GGCCTTTTGTAAAGAAAAATAAAGATGAAGTGTGAAA

[illegible]

Important features:

Signal peptide:
amino acids 1-23

Transmembrane domain:
amino acids 579-599

EGF-like domain cysteine pattern signature.
amino acids 430-442

Leucine zipper pattern.
amino acids 197-219, 269-291

N-glycosylation sites.
amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.
amino acids 124-131, 337-345

N-myristoylation sites.
amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLSKTTTLTSDESVDHTTAGRVVAGQIFLDSESEL
 ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHFPLFLDK
 EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEEAAKRRQMQAEMMYQTGMKILNGSNKKSQKR
 EAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
 SSQAKALVYYTFGALGGNLIAHMLVLSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAGCCCATCTGCACTGGAATTTTCATGAAC TAGCAAGAGGACACCATCTTCTT
 GTATTATACAAGAAAGGAGTGACCTATCACACACAGGGGAAAAATGCTCTTTTGGGTGCTAGG
 CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAAC TAAAGATTGAAGACATCACTG
 ATAAGTACATTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAACTTTTGAT
 AAAAAGGGATTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA
 AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
 CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGTT
 CCGGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
 CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG
 TTATTAATGTCTCCAGTGTTGGAGGTCGCCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA
 TATGCACTGGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC
 ATGCATTGAACCAGGATTGTTCAAAACAAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
 TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAAA
 AGTCTAGACAAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG
 CATGGACCACGCTCTAACAAGTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAA
 TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA
 GCAGAGCTGGCTAATCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGA
 AATTGGCCGATTTCAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT
 CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCAGGGT
 CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCTGCCCT
 GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAATTGAAAGACTTGCCCATTCAAAATGA
 TCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAAC TTGTGAATGTTAAGT
 ATCATCTCTTATCTAAATATTTAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAA

FIGURE 20

MLFWVLGILLICGFLWTRKGLKIEDITDKYIFITGCDSGFGNLAARTFDKKG FHVIAACLTESG
 STALKAETSERLRTVLLDVTDPENVKRTAQWVKNOVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
 REPIEVNLFGLISVTILNMLPLVKKAQGRVINSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
 AFGVHVSCIEPGLFKNLADPVKIEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKG NKS YVNMD
 LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG
 CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTGGGGAAGTAAAA
 GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA
 CATTGAGAAATATATTCCATGCTATCAGCTTTTATAGCTTTTATAATTCCTCAGGCGAAGTAAATG
 AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCGGT
 CGTCATTGAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACTTGCAGGAGCATT
 TTCAAACCAAGACCTTGTTTTCTGCTATTAAACCAAGTATAATAACAGAAAGCTGCTCTACTC
 ATCGACTGGAACATTCTTATATAAACCTCAAAAAGGACTTTTTCACAGGGTACCTTTAGTGTT
 GCCAATCTGGGCATGCTGAACAACCTGGGTTATAAAACTGTATCAGGTTCTGTATGTCCACTGG
 TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTTGAAGAAGATGGATCCTTAAAGGAGG
 TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAAGTG
 GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAACGAGAAATGA
 GAAAAGGAGAGGAGCACAGATTCAGGCAGCAAGAGAGAAGAATCCAAAAAGACCCTCAGGAGA
 ACATTTTCTTTGTGTCAGGCATTACGGACCTTTTTTCCAAATCTGAATTTCTTCATTGATGTT
 ATGTCTTTAAAAATAGACATGTTTCTAAAAGTAGCTGTAACACAACCACCATCTCGATGTAGT
 AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCAC
 AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAGAGATCTCGGTTGTTA
 GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAT
 GAGCAGCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC
 CTACATTTTGATCCTTTAAACCTTACAAGGAGATTTTTTATTTGGCTGATGGGTAAAGCCAAAC
 ATTTCTATTGTTTTTACTATGTTGAGCTACTTGACAGTAAGTTCATTTGTTTTTACTATGTTTACC
 TGTTTGCAGTAATACACAGATAACTCTTAGTGCAATTTACTTCACAAAGTACTTTTCAAACATCA
 GATGCTTTTATTTCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
 ACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
 TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCTGGGCAACGTATT
 GAGACCATGTCTATTAAAAAATAAATGAAAAGCAAGAATAGCCTTATTTTCAAATATGGAAA
 GAAATTTATATGAAAATTTATCTGAGTCATTAAATTTCTCCTTAAGTGATACTTTTTTAGAAGTA
 CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT
 AAAATTTAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGGEVKGEAKNSITDSQMDDVEVVYITIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLDLRWQFKRSRLDQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCCAA
GCAGCGCGCAGCGAACGCCCCGCGCGCCACACCCTCTGCGGTCCCGCGGCGCCTGCCACCCTTCCCTCCTTCCCC
GCGTCCCCGCTCGCCGCCAGTCAGCTTGCCGGGTTGCGCTGCCCCGCGAAACCCGAGGTACCAGCCCGCGCCTCT
GCTTCCCTGGGCGCGCGCCCTCCACGCCCTCCTTCTCCCTGGCCCGCGCCTGGCACCAGGGGACCGTTGCCTGA
CGCGAGGCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCTGCTCGCCTCTTCCACCAACTCCAACCTCCTTCCCC
TCCAGTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGCGCGCTTCCCGTCCGGTCCCAA
GGTGGGAACGCGTCCGCCCCGCGCCACCATGGCAGCGTTGCGCTTGCCCGCGCTTCTTGCACCCTGGCAGTGCTC
AGCGCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGGCAGCTTTTACGTGTCCAAAGGCTTC
AACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGTCCCCAGGGTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAAAGTGTGGTCAGCGAACAGTGCAATCATTTG
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTGATGAATCTTCAAAGAACTACTTGAAATGCAGAGAAATCCCTG
AATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAAATCTGAGCTATTTAAAGATCTCTTCTGATAGTTG
AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG
TTCCGCTGGTGAACCTCCAGTACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG
CCCTTCGGAGATGTCCTCGCAAATGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCGCTACTTTTCGCTCAAGGC
TTAGCGGTTGCGGGAGATGTCGTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCCACTGCCGGGTCTCGTGACTGTGAAGCCATGTTACAACCTACTGCTCAAACATCATGAGAGGC
TGTTTGGCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCTTTCAACATTGAATCGGTGATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCGAGAT
AATAGTGTTCAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTTCTCGT
TCCATCTCTGAAAGTGCCCTTCAGTGCTCGCTTCAGACCACATCACCCGAGGAACGCCAACCACAGCAGCTGGCACT
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTCTGGTCTCCCTTCCGAGCAAC
GTTTGCACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC
CTGTTTGCAGTGACAGGAAATGGATTAGCCAACCAAGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAAACAGAC
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGTGGTGTCCGCTCCTGGGGCA
CAGGCCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGGTTATGCAGAGAGAGTGGAGATTAATTCTCAAACCTCTGAG
AAAAAGTGTTCATCAAAAGTTAAAGGCACCAAGTTATCACTTTTCTACCATCCTAGTGACTTTGCTTTTAAATGAA
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAGTGTGACTTTGTTTCTCATTAGTTTGGG
AGGAAAAGGGACTGTGATTGAGTTGGTTCCCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGATTGTTGATTTTATCACTCTATTATTGTTTGTATGTTTTTTCTCATTTCGTTGTGGGTT
TTTTTTTCCAAGTGTGATCTCGCCTTGTCTTACAAGCAAACAGGGTCCCTTCTTGGCAGTAAACATGTACGTATT
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTATCATGTTATCTTATTAAGAAAAAGCCCAAAAGC

FIGURE 24

MARFGLPALLCTI AVL SAALLAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
 CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH
 LYMQNSSELFKDLFVELKRYVVGNVNLEMLNDFWARLLERMFRLVNSQYHFTDEYLECVSKYTE
 QLKPFQDVPRKLLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL
 VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
 MQDNSVQVSQKVFQCGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
 EKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNEVQVDT
 KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
 KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:**Signal peptide:**

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
 CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
 TACAGCTGCACCGACAGTTGCGAATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
 TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC
 CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCCCTGAG
 AGCCCCGAGAAGAAAATTTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCTGTGATCATT
 TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
 GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG
 AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
 TTCTTCTCCCACTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
 TGTTTTTCAAGATCATTTTGTGTTGCTCTCTCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCT
 GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAACTGTAGCTTCCT
 AGCTAGTGTCAATTAACTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT
 AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLLPLMLSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPCDHFKGNVKKTRHQRHHRKPNKHSRACQQLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
 AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCC
 CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCCTGTGCAAAGCTACTTCCCCCTACCTGA
 TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
 ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
 CAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA
 AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
 GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGCTGTG
 CTCTGTGCAGAGCCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
 TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTC
 GAGCCACCTGGAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
 GAACGCCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
 GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCC
 TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTA
GCAGAATGAGAGAAGACATTATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
 AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
 AAACACTAGGACCCCTGTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCAACGTTTGCCTC
 CCAATGTTGTCCCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCTCCTGAGGCTACAC
 CCATGCGTCTCTAGGAAGTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
 CCTCTCTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG
 ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAA
 ACCACG

FIGURE 28

MDILVPLLQLLVLLLTLPPLHMLLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSSQIKGL
 TGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
 RQLADGSMDDVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQQVFPEPTW
 KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHEMKGAVKQSFSSKALICSFPSL
 QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
 GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCTCTTCTG
 CCACTGACGCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCTTTCTGTGTCTCCTGCCTCAT
 CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTA
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAAA
 GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTCATTAAGAAATT
 GTTCCCTTCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC
 AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAAGTT
 CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
 ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAATCATGTCGG
 GAAGAGATAACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC
 TTCAGATTTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGGTTATTTTGGGATTGTTGTT
 TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
 CAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG
 CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC
 AAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA
 TTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAAGTGCAGGAGCTGCCAGGTT
 ATGGAAGGCGGCCAAGTGGAATATAAGCCCTTTCGGGCATTTCGGTACATGTGGTTCGTACCATTT
 AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAG
 TGGTTACTTGTTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTCTGCTCTC
 TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTAAATCTCTGTGGTGAG
 GATTCCGAGAATCATTTGTCATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
 CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
 AACCAGAATGCATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
 ATTCAAAATCTTGTTCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA
 TTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC
 AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC
 CCATAGTTTTTTATCTGTGTTTGAACTGTGCTGGATGCACTTTCTGTGTTTTGCTGTTGATC
 TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
 AAAAGGAGCAACAAATTAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGAATGAGGA
 GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAAACATT
 TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCCTAGAGAAAAGTTAGTGAATTTTTTT
 TTAAGACCTAATAAACCTATTCTTCCTCAAAA

FIGURE 32

MSGRTILGLCILALALSLAMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYDNDLSIE
 LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTF
 ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQOMTIA
 GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQNALKEQQHG
 ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFD
 FIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
 VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCTTCCTTAGA
ATAATTTGTATGGGATTGTGTATGCGAGGAAAGCCTAAGGGAAAAAGAATATTCATTCTGTGTGGT
GAAAATTTTTTGAAGAAAAAATTCCTTCTTCAACAAGGGTGTCTATTCTGATATTTATGAGGAC
TGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCTTGTGTTTGGCTGGTGACTGGAGTAC
ATTCAAAACAAAGAAACGGCAAAGAAGATTAAAAGGCCAAGTTCACTGTGCTTCAGATCAACTGC
GATGTCAAAGCCGGAAAGATCATCGATCCTGAGTTCATTGTGAAATGTCCAGCAGGATGCCAAGA
CCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGTGGCGCTGCCG
TACACAGTGGTGTGCTTGATAATTCAAGGAGGAAAATACTTGTTCGGAAGGTTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT
TATCGTCTTAGAAAAGTAAACCCAAAAAGGGTGTAACTTACCCATCAGCTCTTACATACTCATCAT
CGAAAAGTCCAGCTGCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACTGCACAGCCGTCCTCTGTATGCAGCTTCTGGCTGTCTAGCTGTGGCCACCCCCAC
CACCTTGCCAAGGCCATCCCTTCTGCTGCTTCTACCACAGCATCCCCAGACCACAATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAACAGGGCC
AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCTTCCAGAACTGTTGGAGC
GGATGTCAGCCTGGGACTTGTTCCAAAAGAAGAATTGAGCACACAGTCTTTGGAGCCAGTATCCC
TGGGAGATCCAACTGCAAAATTGACTTGTCGTTTTTAATTGATGGGAGCACCAGCATTGGCAAA
CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGTATGTTGCCAAGCTCTTGACATTGGCCCTGC
CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAAACCCTGCTACTCACTTTAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAAT
GTAGGTGGGGCCATCTCCTTTGTGACCAAGAATCTTTTCCAAAGCCAATGGAACAGAAAGCGG
GGCTCCCAATGTGGTGGTGGTGTATGGTGGATGGCTGGGCCACGGACAAAGTGGAGGAGGCTTCAA
GACTTGGAGAGAGTCAGGAATCAACATTTTCTTATCACCATTGAAGGTGCTGCTGAAAATGAG
AAGCAGTATGTGGTGGAGCCCACTTTGCAACAAGGCCGTGTGCAGAACAAACGGCTTCTACTC
GCTCCACGTGCAGAGCTGGTTTGGCTCCACAAGACCTGCAGCCTCTGGTGAAGCGGGTCTGCG
ACACTGACCGCCTGGCTGCAGCAAGACCTGCTTGAACCTCGGCTGACATTGGCTTCGTATCGAC
GGCTCCAGCAGTGTGGGGACGGGCACTTCCGCACCGTCTCCAGTTTGTGACCAACCTCACCAA
AGAGTTTGAGATTTCCGACACGGACACGCGCATCGGGGCCGTGCAGTACACCTACGAACAGCGGC
TGGAGTTTGGGTTCCGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC
TACTGGAGTGGTGGCACCAGCACGGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTCTTCAAGAA
GTCCAAGCCCAACAAGAGGAAGTTAATGATCCTCATCACCGACGGGAGGTCTACGACGACGTCC
GGATCCCAAGCCATGGCTGCCATCTGAAGGGAGTGATCACCTATGCGATAGGCGTTGCTGGGCT
GCCCAAGAGGAGCTAGAAGTCATTGGCACTCACCCCGCCAGAGACCACTCCTTCTTGTGGACGA
GTTTGACAACCTCCATCAGTATGTCCCAGGATCATCCAGAACATTTGTACAGAGTTCAACTCAC
AGCCTCGGAACGTAATTCAGAGCAGGCAGAGCACCAGCAAGTGTGCTTTACTAAGTACGTTGT
GGACCACCCACCGCTTAATGGGGCAGCAGCGTGCATCAAGTCTTGGGCAGGGCATGGAGAAAC
AAATGTCTTGTTATTATTTCTTTGCCATCATGCTTTTTCATATTCCAAAACCTTGAGTTACAAAGA
TGATCACAAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGTGGAGATTTTACAT
TTTGACAATTGTTTTCAAAATAAATGTTCCGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG
AGCTTTTGTGAGATTTTAAAGTTGTTATTTCTGATTTGAACCTCTGTAACCTCAGCAAGTTTCAT
TTTTGTCATGACAATGTAGGAATTGCTGAATTAATGTTTGAAGGATGAAAAAATAAAAAA
AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA
AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYNGVQSLSLPRWR
ESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPPIPGTTAQPVTLMQLLAVTVAVA
TPTTLPRPSPSAASTTSSIPRPQSVGHRSEQEMDLWSTATYTTSSQNRPRADPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLPEVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATFNKTHNTSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLPVLRVCDTDLRLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
LTKEFEISDTDTRIGAVQYTYEQRLFEFGFDKYSSKPDILNAIKRVGYWGGTSTGAAINFALEQL
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAAWAAQEELEVIATHPARDHSSF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGCTTTTTTCTGTTTCAGAGTAGTTCACAACAGATCTGAGTGTTTTAATTAAGCATGGAAT
ACAGAAAACAACAAAAAATTAAGCTTTAATTTTCATCTGGAATTCACAGTTTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTATCACGTGGTGCTCTCCGACTACTCACCCGAGTGTA
AAGAACCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGATGGCTCGGCTCTCTGGACTGTCTTCCGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTAC
CTCAGCCTTCCCCACTACAATGTGATAGAACCGCTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTA
CAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAATCCATTTCTGGTCATTC
TGGTGACCTCCACCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCT
TGGTGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGC
ATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA
ACCTGACCTTGAAAACATTATGGCATTGAGGTGGGTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
AAGACAGACACTGATGTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA
GAAGTTTTTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAACCCATATTT
CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTGGGTGGTGGTGGTGGTGGT
TTGGTGCCAAGGATCTATGAAATGATGGGTGACGTAAAACCCATCAAGTTTGAAGATGTTTATGTCGGGAT
CTGTTTGAATTTATTAAGTGAACATTCAATTTCCAGAAGACACAAATCTTTCTTTCTATATAGAATCC
ATTTGGATGCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTT
TGGCAGGTCATGCTAAGGAACACCACATGCCATTATTAAGTTCACATTCTACAAAAGCCTAGAAGGACAG
GATACCTTGTGGAAAGTGTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTGAGTGTGCTGGCTT
ACACTGAACTGAAACTCATGAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTATTAGTCAGG
CCCTTCAAAGATGATATGTGGAGGAATTAATATAAAGGAATTGGAGGTTTTTGCTAAAAGAAATTAATAGG
ACCAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGTAAAACAAAACAATGTAGAGTTTATTTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA
TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAACCTCTTCACTGAAGTTATA
CTGAACAAAATTTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT
ATTATTTAAAATTACTTCACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAG
TGAATCATTCTTTACATGCAACATTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCACTCCA
TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTGGACTTTGTTAAATATTTTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEAEEKDKMLA
LSLEDEHLLYGDIIRQDFLDTYNNLTTLKTIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNHIPEDTNLFYRIHLDVCQLRRVIAAHGFSSKEITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCGCGGCAAGGATGGAGCTGGGTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCCTTCTCATC
TCGTCTTGGCAAGAGAGTACACAGTCATTAAATGAAGCCTGCCCTGGAGCAGAGTGAATATCATGTGTCGGGAGTGTGTG
AATATGATCAGATTGAGTGCCTGCTGCCCCGAAAGAGGGAAGTCGTGGGTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAAACTGCAAGAGCTGCCGAAATGGCTCATGGGGGGGT
ACCTTGAGTACTTCTATGTGAAGGGGTTCTACTGTGCAGAGTCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTTGTGCGAGCCCCAAGGGTCAGATTTGTGGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTCATGC
TAAACCTGGGTTTGTTCATCCAACCTAAGATTGTTCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTTCGTGATGGAGACAACCGCATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCTCACTCCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGC
ATGCTCCTCATCCCCCTTGTTCATGACGGCAGTGCCTCTTGACAAGGCTGGATCTTACAAGTGTGCTGCTGGCAGGC
TATATCGGGCAGCGCTGTGAAAATCTCCTTGAAGAAAGAACTGCTCAGACCTTGGGGGCCAGTCAATGGGTACAGAAAA
TAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATGGCACCGTGGTGTCTTTCTTTTGAACAACCTCTATGT
TCTTAGTGCATGAGAAAGAACTTGGCAGCAGAAATGGAGAGTGGTCAGGGAAACAGCCCATCTGCATAAAAGCCTGCCGA
GAACCAAAGATTTCAGACCTGGTGAGAAGGAGATTCTTCCGATGCAGGTTCACTCAAGGGAGACACCATTACACCAGCTAT
ACTCAGCGGCTTCAGCAAGCAGAAACTGCAGAGTGGCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGGCGGGCACCATCTGCATCCCTATCTGCGGAAAATTGAGAACATCACTGCTCCAAAGA
CCCAAGGGTTGCGCTGCGCTGGCAGGAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCTTACACAAGGGAGC
GTGGTTCTTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCATGTGTTACTGACCTGGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTGGGGAAATCTACCGGGATGATGACCGGGATGAGAAGACCATCC
AGAGCCTACAGATTTCTGCTATCATTTCTGCATCCCAACTATGACCCCATCTGCTTGATGCTGACATCGCCATCCTGAAGCT
CCTAGACAAGGCCCGTATCAGCACCCGAGTCCAGCCCATCTGCCCTCGCTGCCAGTCCGGGATCTCAGCACTTCTTCCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGTCTTGGCAGAGTGGAGAGCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGG
TGGTCAGTGTGGTGGACTCGCTGTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCTTCTGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTGCTCTCCCGGGA
CGAGCATCTCCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCCTTACCAAGGTGCTGCCTTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAG
TGTTTCTGTATATCCGTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTGGCCTGTGAAGTGTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTGAAGTACCTCCATTGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTTCTTCAAGAAGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGGTCTTCCCAACTTTCAGTTATACGAATGCCATCAGCTTGACAGGGAAGATCTGGGCTTCATGAG
GCCCCCTTTGAGGCTCTCAAGTTCTAGAGAGCTGCCTGTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCCT
TGTTGATCATGGCCACAGTACAGTCTGGTCTTTTCCCTCCCATCTCTTGTACACATTTTAAATAAAATAAGGGTTGGCTTCT
GAACTACAAAAA
AA
AA

FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTATTAAACGTGGCTTAATC
 TGAAGGTTCTCAGTCAAATCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG
 TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAAATGAAGGCGCTTCTGTTGC
 TGGTCTTGCCTTGGCTCAGTCTGTAACTACATTGACAATGTGGGCAACCTGCACCTTCTGTATTAGAACTCTGTA
 AAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC
 TCACAGCCACGGCTCCCTCCCAAGAGGTTTCTGCAGCTGCCACCATCTCTTAATGACAGACGAGCCTGGCCTAGACA
 ACCCTGCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAATA
 GGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTCTTCGAA
 GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAAATTCTGAAAACACCACTGCCCTGAAG
 TCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAAATTACCAGCATCAAGATCAATCGAGTAGATCCAGTGAAA
 GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCACTGGTCCATATCATTATCCAACACATTTATCGTGATGGGG
 TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
 ACAAACACGCTGTGCGTCTCCTGCGGCAGCCCTGCCAGGTGTGTGGCTGACTGTGATGCGTGAACAGAAAGTCCGGA
 GCAGGAACAATGGACAGGCCCGGATGCCTACAGACCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC
 CCGAGGACGAGCTTGAATAAACTGGTGCSCAAGGTGGATGAGCCTGGGGTTTTCATCTTCAATGTGCTGGATGGCG
 GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTAGCCATCAATGGACATGATCTTCGATATGGCA
 GCCCAGAAAGTGGGCTCATCTGATTGAGGCCAGTGAAGACGTGTTACCTCGTCTGTCGCCGCGAGGTTCCGCGAGC
 GGAGCCCTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCAGGGCCAGGGGAGAGGAGCAACA
 CTCCCAAGCCCTCCATCTACATTAATTAATGTCATGAGAAGGTGGTAAATATCCAAAAAGACCCCGGTGAATCTCTCG
 GCATGACCGTCGCAGGGGGAGCATCACATAGAGAATGGGATTGCTATCTATGTCATCAGTGTGAGCCCGGAGGAG
 TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAACTGACAGAGGTGAGCC
 GGAGTGAGGCGAGTGGCATTATTGAAAAGAACATCATCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC
 CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCCATCCT
 GGGTCATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTACGAAGAAACAGCTGGAAGTC
 TGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAACAAACCTTTTTTCATCAATCCATTGTTGAAGGAA
 CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTGCTGTCAATGGTAGAAGTACATCAGGAATGA
 TACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT
 TTTTATAGAATCAATGATGGGTCAGAGGAAAACAGAAAAATCACAATAGGCTAAGAAGTTGAAACACTATATTATC
 TTGTGAGTTTTATATTAAAGAAAGAATACATTGTAAAAATGTCAGGAAAAGTATGATCATTAATGAAAGCCAGTT
 ACACCTCAGAAAATATGATTCCAAAAAATTAATACTACTAGTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTAC
 AACATTGTTTATATTTTCTATTCAATAAAAGCCCTAAAACAATAAATGATTGATTGTATACCCCACTGAATT
 CAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTAAATTTACAGCT
 AAAATATTTTAAATGCAATTGCTGAGAAACGTTGCTTTCATCAACAAGAATAAATATTTTTCAGAAGTTAA

FIGURE 40

MKALLLLVLPWLS PANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGC PDGCASLTATAPS
 PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVD SGRSNRTRARPFERSTIRSRSFKKINR
 ALSVLRRTKSGSAVANHADQGRESENTTAP EVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
 GGSETPLVHII IQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
 REQKFRSRNNGQAPDAYRPRDDSFHVLNKSSPEEQ LGIKLVRKVDEPGVFI FNVLDGGVAYRHG
 QLEENDRVLAINGHDLRYGSPESA AHLIQASERRVHLVVS RQVRQ RSPDIFQEAGWNSNGSWSPG
 PGERSENTPKPLHPTITCHEKV VNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISR DGR
 IKTGDI LLNV DGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAP P
 SDWSPSWVMWLELP RCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
 RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
 467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
 CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGACT
 TTCCACCTTTCCCTACAAATCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGTTGGGTGGGC
 CACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCTTAAAGCAAAGGAGTTCATGGCTAATTTCC
 ATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAAGTT
 GACAACTGTCCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACAGATCTCAC
 TTTGGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAG
 CTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
 CATCTGCATCCCTTCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGG
 TAAAAAGTTTAAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
 ACTGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG
 CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTGG
 GGGTGTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACACTGCGGGAT
 GGGGAGGCCAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG
 CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
 GATGAAGCTCTTACACCAAGTGTACAGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAAT
 TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTTCTGGTTTGGTGCATTGACCC
 TGGATCTTTTGGTGATGTTTGAAGAAGTGAATCTTTGTTTGAATAATTTTGGCCTAGAGACTTCAA
 ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTGTATTTTCT
 TAGCAGAGCTCCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGAT
 CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
 AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGAT
 AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
 AGGTACGAAGATACAATACTGTTATTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTGAGGT
 GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAAGTTGGGAATGAAGA
 GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
 CTTCAGGGGAGGACCTGCCAGGTATGCCTTCCAGTGATGCCACCAGAGAATACATTCTCTATTAGT
 TTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
 ATTAACATAATAAATATGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

FIGURE 42

MGFNLT FHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNGFSNNYWGWGGE
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 44

MALSSQIWAACLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCAACATGCCTCACCC
 TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
 TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
 GACCTTCAACACAACCCCTCTTGTACCATAACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
 ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAG
 AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
 GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
 ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT
 TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
 CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCGTGCAGCAGAACT
 TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
 GTCCTCCTGTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG
 GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTGCGG
 AAACCTCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
 AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
 GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA
 TCTAGACAGCAGTGCCTCCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIVTQNRNRERVDFFDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTHEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTTCTCAAAACCCCATCTCTTGCTTTGAGTGGTGGTTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA
CTTGCACTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTTCTCAGTATTTTTAGGTCTATTGCTTGTGGAATTCTGGAGGTCCTGTTTGGGCT
CAGTCAGATAGTCATCGGTTTTCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKKRACCNRTGMFLSSFFSVITVIGALYCLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSSEENKHRLIHFSVFLGL
LLVGILEVLFLGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
 AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
 ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA
CATGGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAGCCAATGACC
 CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
 TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
 CAGCCAGAAGCAGCACAGTCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
 CTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
 TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCCCT
 TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACCCCTCA

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
 AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
 AGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
 GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGGCCCCCTGCAGAG
 CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA
 GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGCAGCTGGCTCTAAAGTCAGTGAG
 GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
 AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACTTGGGCACGAGA
 TTGGCAGACAGGCAGAAGATGTCATTTCGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGG
 GTGCCTGGCCACAGTGGTGCTTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
 CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG
 GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA
 GGGCCACCAAACCTTTGGGACCAACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCAAGTGAG
 AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
 ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
 AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
 TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
 CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGA
 CATAAACCCGGGTGTGAAAAGCCAGGGAATGAAGCCCGCGGAGCGGGGAATCTGGGATTCAGGG
 CTTCAAGGACAGGGAGTTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATCGCCTCCTTG
 GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
 GGTGGAGTCAATACTGTGAACCTCTGAGACGTCCTTGGGATGTTTAACTTTGACACTTTCTGGAA
 GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC
 GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACTCCCTCCTTAA
 AACACCACCCTCTCATCTACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAA
 AA
 AA

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2} dt = 1$$

MKFQGGLACLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIKEAGGAAGSKVS
 EALGQGTREAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVF
 GHSGAWETSGGHGIFGFSQGGGLGGQGGNPGGLTPTWVHGYPGNSAGSFGMNPQGAWPQGQGGNGGPPNF
 GTNTQGAVAPQGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGSGSGSQSGSSGSGSNGDNNNGSSSSGGS
 SSCSSSGSSSGSSGGSSGGSSGNSGSGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKPGCEKPGNE
 ARGSGESGIQGFRCQGVSSNMREISKEGNRLGGSGDNYRGQSSWGSGGGDVAGGVNTVNSETPSGM
 FNFDTFWKNFKSKLGFINWDAINKDORSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPKRNWFWGHLG
 LITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
 WLGEILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
 SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
 RLVHDFTDVIRERRRTLPTQGIDDFKDKAKSKTLDLDFIDVLLLSKDEDGKALSDEDIRAEADTF
 MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
 LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
 LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
 VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA
 GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT
 TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCACTCTTTGGCATTGACGTGG
 TACAGCCTTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCGGTGTGTCTTGC
 ATAATTTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACATATGGACAGAAGCTGGTGGACAGTTTT
 GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGCAGCAATGTGTTGCTT
 GTGATTGCAACATTTGAGGGTTACTTTTGGGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
 GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGAATCTTCCTCATGTACCTGTTTCCTC
 TCTGGATGTTGTCCCACTGAATCCCATGAATACAAACCTATTCAGCAACAGCAAAAAAAAAAAAA
 AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGAGTGGAGCACCAGCAGGCCGCCAACATGCTCTGTCTGTGCCTG
TACGTGCCGGTCATCGGGGAAGCCAGACCGAGTTCCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC
ATTTTCAAGCTCAGTGTCTTTCATCCCTCCAGGAATTTCCACCTACCGCCAGTGGGAAGCAGAAAATTTGACAAGCT
GGAGATAAGGACCTTGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGATTTTGGACAAAAAGAAATGATGGACGCATTGACGCGCAGGAGATCATGAGTCCCTGCGGGACTTG
GGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTTCTCAAGAGCATGGATAAAAACGGCACGATGACCATCGACTGG
AACAGTGGAGAGACTACCACTCTCTCCACCCGTTGGAAAACATCCCCGAGATCATCTCTACTGGAAGCATTCCACG
ATCTTTGATGTGGGTGAGAACTAACGGTCCCGGATGAGTTACAGTGGAGGAGAGGCAGACGGGGATGTGGTGGAGA
CACCTGGTGGCAGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTGAGAAGGAGGGGCCAGG
TCACTCTGGCGGGCAATGGCATCAACGTCCTCAAAATTTGCCCGCAATCAGCCATCAAAATTCATGGCCTATGAGCAG
ATCAAGCGCCTTGTGGTAGTGACCAAGGAGACTCTGAGGATTACGAGAGGCTTGTGGCAGGGTCCCTGGCAGGGGCC
ATCGCCAGAGCAGCATCTACCAATGGAGGTCTGAAGACCCGGATGGCGCTGCGGAAGACAGGCCAGTACTCAGGA
ATGCTGGACTGGCCAGGAGGATCTGGCCAGAGAGGGGTGGCCGCCTTCTACAAAGGCTATGTCCCAACATGCTG
GGCATCATCCCTATGCCGGCATCGACCTTGCAGTCTACGAGACGCTCAAGAAATGCCTGGCTGCAGCACTATGCACTG
AACAGCGCGGACCCGGCGTGTGTGCTCTGGCCTGTGGCACCATTGTCAGTACCTGTGGCCAGCTGGCCAGCTAC
CCCCTGGCCCTAGTCAGGACCCGGATGACGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCTC
TTCAAAACATATCCTGCGGACCGAGGGGCCCTTCGGGCTGTACAGGGGGCTGGCCCCCAACTTCATGAAGGTATCCCA
GCTGTGAGCATCAGTACGTGGTCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTCCGGGTGACGGGGGAGGGC
CGCCCGGAGTGGACTCGCTGATCTGGGCGCAGCTGGGGTGTGCAGCCATCTATTCTGTGAATGTGCCAACACT
AAGCTGTCTCGAGCCAAGCTGTGAAAACCTAGACGCACCCGAGGGAGGGTGGGGAGAGCTGGCAGGCCAGGGCTT
GTCTGTGACCCAGCAGACCTCTGTGGTTCCAGCGAAGACACAGGCATTCTTAGGGTCCAGGGTCAGCAGG
CTCCGGGCTCACATGTGTAAGGACAGGACATTTCTGCAAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGGCT
TAGTTCTTCCATTTACCCCTGCAGCCAGCTGTGGCCACGGCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGC
CCTCTTGTGCTGCCTGTCTGTCTGAGGTAAGGTGGGAGGAGGGTACAGCCACATCCCACCCCTCGTCCAATCCC
ATAATCCATGATGAAAGGTGAGGTACGTGGCCTCCAGGGCCTGACTTCCCAACCTACAGCATTGACGCCAACTTGGC
TGTGAAGGAAGAGGAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATG
CTTGGGAGTGCAGGGGGCTCGGGCTGCCTGGCCTGGCTGCACAGAAGGCAAGTGTGGGGCTCATGGTGTCTGAGCT
GGCCTGGACCTGTGAGGATGGGCCCCACCTCAGAACCAAACTCACTGTCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTTGAAGGCGAAGGGCAGAGCGTTTGTGTCTGGGAGGGAAGGAAAGGTGTTGGAGGCCTTAATTATGG
ACTGTTGGGAAAAGGGTTTGTCCAGAAGGACAAGCCGACAAATGAGCGACTTCTGTGCTCCAGAGGAAGACGAGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGTTCCTGTCCAACCCAGCAGGGGGCGCAGC
GGGACAGCCCCACATTCACCTTGTGTCACTGCTTGGAACTATTATTATTATTATTGAACAGAGTTATGTCTCT
AACTATTTTATAGATTTGTTAATTAATAGCTTGTCAATTTCAAGTTCAATTTTATTCATATTATGTTCTATGGTT
GATTGTACCTTCCCAAGCCCGCCAGTGGGATGGGAGGAGGAGGAGGAGGGGGCTTGGGCGCTGCAGTCACATCT
GTCCAGAGAAATTCCTTTGGGACTGGAGGCAGAAAAAGCGCCAGAAAGCAGCAGCCCTGGCTCCTTTCTTTGGCAG
GTTGGGAAGGGCTTGCCCCAGCCTTAGGATTTAGGGTTTGAAGTGGGGGCTGGAGAGAGAGGGAGGAACCTCAAT
AACCTTGAAGGTGGAATCCAGTTATTTCTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTCTGAATGTCAAGGCAG
TGAGGTGCCTCTCACTGTGAATTTGTGGTGGGCGGGGCTGGAGGAGAGGGTGGGGGCTGGCTCCGTCCCTCCCAGC
CTTCTGCTGCCCTTGTAAACATGCCGGCCAACTGGCGACCTCACGGTTGCACTTCCATTCCACCAGAAATGACCTGA
TGAGGAAATCTTCAATAGGATGCAAGATCAATGCAAAAATTTGTTATATATGAACATATACTGGAGTGTCAAAAAAG
CAAAATTAAGAAAGAAATGGACGTTAGAGTTGTCAATTTAAAGCAGCCTTCAATAAAGTTGTTTCAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAFQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKDLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIIMQSLRDLGVKISEQQAEEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTPDEFTVEERQTGMWWRHVLVAGGG
AGAVSRTCTAPLDRKVLMMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIAK
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQSSIIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNWALQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACCTTTGAACCTGACATCAAACCTTCTGATAT
CGTGATAACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTGGAGCAGGATGAAATGTTTCAGAGGCCGACAGCAGTGTGCTGATCAAGTGATAGTT
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGGCGGTGTGAGGCTCCCCGATGGTTCCCC
CAGCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACTTCTCGGAAGTCTCCAATAC
CAGCTTTGAGCTGAACCTCTGAGAATGTGACCATGAAGTTGTGTCTGTGCTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
ACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT
CTCTTCTTTCTTTGCCATCAGCTGGGCACCTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAT
GTGCCTTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTAC
CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTGATGTGAAGTACAGTGTGTTAAGA
GTGATAAGTAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATCCAC
AAATTAAGCTGTAGTATGTACCTTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG
GCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCTTGCTTC
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTTAAACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTSEIKRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDS DH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
 TCCCAGGCCTCTTTGCCCCGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACTTCGGGACC
 AACTTGCCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCAGCCCGC
 TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCTCTGAAGCTCAGCGTGCCCTCCATCAGATG
 GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
 GATTCTTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCTGGGGGA
 AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCGGCCCTCGCTCCGGGCAGTGGCCCTT
 TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
 TCGGAGTCCAGACGACTGCCCCGTTCTAATTCACTGGGAGCCGGGGGAAAAATCCTTTCCCAACG
 CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCAGTG
 TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCTGAGGGA
 ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG
 CAGCTGGGGAAATATTAATCGGTATCCAGGAGGAGCTGGGGAAATATTAATCGGTATCCAGGAG
 GCAGCTGGGGGAATATTCATCTATACCCAGGTATCAATAACCCATTTCCCTCCTGGAGTTCTCCGC
 CCTCCTGGCTCTTCTTGGAAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
 GGGCTAGAGCACGATAGAGGGAAACCAACATTGGGAGTTAGAGTCCTGCTCCCGCCCCCTTGCTG
 TGTGGGCTCAATCCAGGCCCTGTTAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCCTCCCC
 TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AA

FIGURE 64

MQGRVAGSCAPLGILLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGSPNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
 TGGGCTGCCCCCTTGTCCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
 CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA
 TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC
 AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCTTCTTGGCCC
 GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCACCCCTC
 CTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 66

MSGGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGC
 CAGGTGCCCCGTCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC
 TTCCTCGGGCGCTGCCAACC CGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG
 CTGGGCCTGCCGTTCTTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC
 AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
 AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG
 GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
 GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
 GCCTGCCCATCTAGGTCCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
 GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA
 AGACTCTGCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTTTTATATATTTATATAAAATTAG
 TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLARWGRAWGQIQTTSANENSTVLPSTSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:**Signal peptide:**

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTAGAGGTTTGTTCCTCTAGTTCTGTGCCTGCTGCACCAG
 TCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATATTGTCATTGTTATAGATCCTAGTGTGCCAGAA
 GATGAAAAATAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA
 TTTTTCCTCAAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAAACATGAA
 AACATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCAGGTAGAGATGAACCATACACCAGCAGTTCA
 GAATGTGGAGAGAAAGGCGAATACATTCACTTCACCCCTGACCTTCTACTTGGAAAAAACAAATGAATATGGACCA
 CCAGGCAAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGATGAGTACAATGAAGATCAGCCTTTC
 TACCGTGCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT
 CAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAAATTGATTCTACAACAAAACCTGTATGGAAAAAGATTGCAATTCCTT
 CCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATGCAAAAGTATTGATTCTGTTGTTGAATTTTGTACGAA
 AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAGTGCAATTTTGAAGTACATGGGAGGTGATTAGCAAT
 TCTGAGGATTTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTATTGCTGAAGATCAGTCAA
 AGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAAGGACCGCCTAAATCGAATGAATCAAGCAGCA
 AAACATTTCTGCTGCAGACTGTTGAAAAATGGATCCTGGGTGGGGATGGTTCACTTGTATAGTACTGCCACTATTGTA
 AATAAGCTAATCCAAATAAAAAAGCAGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
 ACTTCCATCTGCTCTGGAATTAATATGCAATTCAGGTGATTGGAGAGCTACATTCCTCACTCGATGGATCCGAAGTA
 CTGCTGCTGACTGATGGGAGGATAACACTGCAAGTTCTTGTATTGATGAAGTGAACAAAGTGGGGCCATTGTTTCAT
 TTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTTATGTT
 TCAGATGAAGCTCAGAACATGGCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAAAG
 TCCCTTCAGCTCGAAAGTAAGGGATTAAACACTGAATAGTAATGCCTGGATGAACGACACTGTCTATAATTGATAGTACA
 GTGGGAAGGACACGTTCTTCTCATCACATGGAACAGTCTGCCTCCAGTATTTCTCTGGGATCCAGTGGAACA
 ATAATGGAAAAATTCACAGTGGATGCAACTTCCAAAAATGGCCTATCTCAGTATCCAGGAACTGCAAAAGGTGGGCACT
 TGGGCATACAATCTTCAAGCCAAAGCGAACCAGAAACATTAATATTACAGTAACCTTCTGAGCAGCAAAATCTTCT
 GTGCCTCCAATCACAGTGAATGCTAAAAATGAATAAGGACGTAAACAGTTTCCCAGCCCAATGATTGTTTACGCGAGAA
 ATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTCATTGAATCACAGAATGGACATACAGAAGTT
 TTGGAACCTTTGGATAATGGTGCAGGCGCTGATTCTTCAAGATGATGGAGTCTACTCCAGGTATTTTACAGCATAT
 ACAGAAAATGGCAGATATAGCTTAAAGTTCCGGGCTCATGGAGGAGCAAACTGCCAGGCTAAAAATACGGCCTCCA
 CTGAATAGAGCCGCTACATACCAGGCTGGGTAGTGAACGGGGAAATGAAGCAAAACCCGCAAGACCTGAAATGAT
 GAGGATACTCAGACCACCTTGGAGGATTTGAGCCGACAGCATCCGGAGGTGCATTTGTGGTATCACAGTCCCAAGC
 CTTCCCTTGCCTGACCAATACCACCAAGTCAAATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTCTT
 ACATGGCAGCACCAGGAGATAATTTGATGTTGGAAAAATTCACAGTTATATCATAAGATAAGTGAAGTATTCTT
 GATCTAAGAGACAGTTTGTATGATGCTCTTCAAGTAAATACTACTGATCTGTCCACAAAGGAGGCCAACTCCAAGGAA
 AGCTTTGCATTTAAACAGAAAAATATCTCAGAAGAAAAATGCAACCCACATATTTATTGCCATTAAAGTATAGATAAA
 AGCAATTTGACATCAAAAGTATCCAACATTGCACAAGTAACCTTGTATCCCTCAAGCAAAATCCTGATGACATTGAT
 CCTACACCTACTCCTACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTTCTACGCTGGTATTG
 TCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAAGTACCACCATTTGAACCTTAACGAAGAAAAAATCTTC
 AAGTAGACCTAGAAGAGAGTTTAAAAAACAAAAACAATGTAAGTAAAGGATATTTCTGAATCTTAAAAATTCATCCCAT
 GTGTGATCATAAACTCATAAAAAATAATTTTAAAGATGTCGGAAAAGGATACCTTGATTAAATAAAAAACACTCATGGATA
 TGTAATAAAGTGTCAAGATTAATAATTTAATAGTTTCATTATTGTTATTGTTGTAAGAAATAGTGATGAACAAAG
 ATCCTTTTTCATACCTGATACCTGGTTGATATTATTTGATGCAACAGTTTCTGAAATGATATTTCAAATTTGCATCAA
 GAAATTAATAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAATTAACCAACATTTGGAAAAAATAA
 AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNTSFIKLNNNGFEDIVIVIDPSVPEDEKIEQIEDMVTASTYLFE
 ATEKRFFFKNVSIIPENWKENPQYKRPKHENHKKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
 IHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN
 RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMFMQSIDSVVEFCNEKTHNQEAP
 SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLKISQRIVCLVLDKSGSMGGKDRLNR
 MNQAAKHFLLOTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSCSGIK
 YAFQVIGELHSQLDGSEVLLLTGDNNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
 SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL
 ITWNSLPPSISLWDPSTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
 AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
 GADSFKNMGVYSRYFTAYTENGRYSLKVRAGGANTARLKLRPPLNRAAYIPGWVNGEIEANPP
 RPEIDEDTQTTLDEFSRTASGGAFVVSQVPSLPLPDQYPPSQITDLATVHEDKIIILTWTAPGDN
 FDVGKVQRYIIRISASILDRLDSFDDALQVNTDLSPKANSKESFAFKPENISEENATHIFIAI
 KSIDKSNLTSKVSNIQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI
 VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW
EQHRNYVSSLKRIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQV
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEATIESALET
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVNEKLN
TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKA
NFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNLCRLNTQPGKKV
FYFVLF
SQYNPGIIYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDLDIKG
WGGE
DVHLYRKYLHSNLIVVRTPVRGLFHLWHEKRCMDELTPQYKMCMSKAMNEASHGQLGML
VFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTFDPDLPALQ
PHGYDQPGTEPLDLHNNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPPGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLGTLFSTEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTC
TCTGGTGGTTTGCCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA
ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCTGACAGCTCC
AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCCTTCGTCCC
AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTCTTTTTT
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAAACCCAGCAAATTT
GATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTA
ACTTTATCACCCCTCAATATCTCGGATGATTTCAAATTTCTCATCAGGATATGAGTTTACTGGGA
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTGTGACTCTGAAG
AAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTGAGAACCCTGACATTTGTGCGGGGCTGAAGAGCAGGA
GCTCAGTTTGAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
TCTTGGGCCCCGAAACGTTACAGTACTCATAACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCTGACTGGGA
TCCCCAACTGGCAGGCTGTGTATTCCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG
AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGATGCCAACACTTCCTTTTGCTTTTGTTCCTGTGCAACAAAGTGAG
TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTCTAGTGT
CTGTGAGAATTACTTATTTCTTTTCTTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA
GGTCTCTTAACAATGATGGTGGGCTCTGGAGTCCAGGGGCTGGCCGGTTGTCTATGCAGAGAA
AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISSVVLTAPEKWKRNPEDLPVSMQQIY
 SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK
 DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEDKRFFVPAEK
 IVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
 DSEENTEGTSLTQQESLSRTIPDKTVIEYEYDVRTTDCAGPEEQELSLQEEVSTQGTLLLESQA
 ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
 EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
CTGGGAAGATGGCCGGCCCGTGACCTTCACCCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
CAAGCCACCCTCAGTCCCCTGAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCTGAAGCACATC
ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC
CCCCCCGCTGGTCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAATGCTGTGA
TAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGCAGGTGATGAACCTCCTAGTGCCATCCCTGC
CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTGA
TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGACAACATCCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
ATTCATGGTCTGTGGACTCTGTGCTTCCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAATCAGGAC
ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTT
TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACTCTGGGATTGGCTGGTTCACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG
CCTTGGGATTCGAGGCAGCTGAGTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGG
GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACAATTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKHDNATSILQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITQLYLGAKLKLLDSQGKVT
KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAABAVALSPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNLISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSVLKALG
FEAAESSLTKDALVLTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACCTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACAGCACAGGCATCACCAGTGTGACATCTATAGC
ACCTTCTGGGCCTGCCCCGTGACATCCAGGCTGCCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTCATCCTTGAGGCCTCCTGGGATTCT
ATCCTGTTGCCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCAGTGGTGCCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCTGATAG
CTGGAATCATCCTCTGCTTTTCTGCTCATCCAGAGAAATCGCTCCAACCTACTACGATGCCTAC
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTGGAAGAACCAGGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC
TCACCTTGCTGCTCCCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
TCCCACTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTGTTATGACTCCACAGTGTCCA
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYYLGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLLPADIQAAQAMMTSSAIISSLACIIISVGMRCCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTT
 CCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCACCTTAGACCTCCCTTCCTGCCCTCCTTTCCCT
 GCCCACCGCTGCTTCTTGGCCCTTCTCCGACCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
 GTTGATCTGTGGCCCTGTGCCTCCGTGTCTTTTCGTCTCCCTTCTCCCGACTCCGCTCCCGG
 ACCAGCGGCTGACCTGGGGAAGGATGGTTCCCGAGGTGAGGGTCTCTCTCCTTGTGTGGGA
 CTCGCGCTGCTCTGGTTCCCCCTGGACTCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT
 CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
 ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCCT
 GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
 TCCCTCTGGACTCCGGGCCCCACCAAGTCTGCCAGCACAACGGGACCATGTACCAACACGGAG
 AGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCTGCCCAACAGTGTGTCTCTGCAGCTGC
 ACAGAGGGCCAGATCTACTGCGGCCTCACAACCTGCCCCGAACAGGCTGCCAGCACCCCTCCC
 ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGCAATCGGATGAAGAGGACA
 GTGTGCAGTCGCTCCATGGGGTGAGACATCTCAGGATCCATGTTCCAGTGTGCTGGGAGAAAG
 AGAGGGCCCGGCACCCAGCCCCACTGGCCTCAGCGCCCTCTGAGCTTATCCCTCGCCACTT
 CAGACCCAAGGGAGCAGGCAGCACAACCTGTCAAGATCGTCCTGAAGGAGAAACATAAGAAAGCCT
 GTGTGCATGGCGGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC
 CCCTTGCCCTGCATCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
 CACCGAGTACCCCTGCCGTCACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTGCCCCAGAGG
 ACAAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGGTC
 CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC
 CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAGATGAGGAACTGAGGCTCAGAGAG
 GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG
 GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGTGGCCCCACGAAGGTCACT
 GGAACGTCTTCTAGCCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG
 ACATAACAAAGACC**TAA**CAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA
 TAAGAAGTTGCATTACCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPVHCPQPVTEPQQCCPKVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTYS
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTTEYPCRHPKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWPFRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 84

MALMLSLVLSLLKLGSGQWQVFGPKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVVH
 LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAIWELQ
 VSALGSVPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPQGDLSTDSRTNRDMHGLFDVEISL
 TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
 FQWKIQAELDWRRKHGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
 TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRNLNGEHLFT
 LNPRFISVFPRTPTTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
 PIVICPVTQESEKEASWQRASAIPESTSSESSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTTCCTGTA
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCTACCCCTCGCATGGCTGGATT
TACCCTGGCCCAGTAGTTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA
TACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCCAGGCACCTGGAGTCCGGCTGCCCCC
AGAACTTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTGAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCTACCCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCT
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCTGTGCCCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTGGGGGAGCTGGAG
CCACAGCCCTGGTCTTCCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTGAGGGGTTT
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCAG
CTTCTGCCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGGACAGGAGGCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC
TCCCTTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWFREG
 ANTQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
 KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVSPLDPSTTRS
 SVLTLPQPQDHGTSITCQVTFFGASVTTNKTVHLNVSYPPQNLMTVTFQGDGTVSTVLGNGSSL
 SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
 LGSQQVYLNVSLSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDGTGIE
 DANAVRGSSASQGPLTEPWAEDSPPDQPPASARSSVGEGLQYASLSFQMKPWDSRGQEATDTE
 YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
CCAAGGAAAGTGCACTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCTAGTGCATTTGA
TGGCCTGTATTTCTCCGCACTGAGAATGGTGTTATCTACCAGACCTTCTGTGACATGACCTCTG
GGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCTTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCTAT
GGCCAGCGGGAATTCAGTGCGGGATTTGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
GGAACATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLFLFIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
 YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
 SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGLIYQKYPVKYG
 EGKCWTDNGPVI PVVYDFGDAQKTASYSPYGQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
 TEHHCIGGGGYFPEAS PQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTGCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACCTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCTCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTCGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG
CCTGCCCTCTCAGTTCACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGAGTCTCCATGTTCAAC
AGCACCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT
ATTGAGCGTTGGGAGTGCCTACGTATTGGTACATTTTGTCTTCTGCAGTGCCCTTCCAGCTG
TCACTGAAATGGCTTTATTTCGTACCGTCTTTGGGCTGAAAAAGAAACCCTTCTCAATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGCAAGGGCCGCTTCGTATTCCCTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLÄHLVVVITLFWSRDSNIQACPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
 TCGCTGCTGCTTCGTGTTCTTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
 GTCTTCCCGGCTCAGAGGACCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
 CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCTTAGGGCTGCT
 GGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCGAACCACAGCCCCC
 CACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
 GCCCTGAACCTGCTCGTTCACAGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
 CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
 AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG
 GAGTGGGAGAAGGTAGAACGGGGCCGCCGGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG
 CTCCCAGAGACCAGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG
 TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTAC
 CATAGTGATAACCCCTACTACCCATCTGGGTGACCCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
 TGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCTC
 AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
 AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
 AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG
 TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
 CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA
 GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
 GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCTTGTCGTGTGCTGAGCATGG
 CATGAGGCTGAAGTGGAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
 CAGGCCACCCCTTTCCAAAATTCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
 GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC
 CGCGTGCTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG
 GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTGACACGGGGTCTCCCGGATCTGGATGGCGC
 CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGCCGAGAGCATGTGCTGGATCTGTTC
 TGTGTGCTGTGCTGTGGGTGGGGGGAGGGGAGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT
 TGTGTGAAGAATCGTGTCTTGGAGCAGGAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPKRGRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 94

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLIAYVSGLGFGIMSGV
FSFVNTLSDSLGPGTVGIHGDSPOFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKCLLCQDKNFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTTLSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

3 GGGAGGGGAGGGTCGG

GCTCAAGTGCCTTGCCTTTGCCCAACCCAGCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
 CTTCTTGTCTTGGCAGCTGGACAAGGGAGCCAGTCTTGGCGCCTGGAGGGCCTGTCTTGACCATG
 GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCAGGCTCTCCCCAAGGCCAGCCAGCTGC
 AGAGCTGTCTGTGAAGTTCCAGAAACATATGTGGTGAATTTCCCTTTATACCTGACCAAGTTGC
 CGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGTCTGTGAGGGGACTCAGGCAAGGCAACTGAG
 GGCCCATTTGCTATGGATCCAGATTCTGGCTTCTGTGTGGTACCAGGGGCTTGGACCGAGAGGA
 GCAGGCAGAGATTACAGCTACAGGTACCCTGGAGATGACAGGATGGACATGTCTTGTGGGTCCAC
 AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACAGGTGCCCATTTCTCTAAGCCATGTAC
 AGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCCTTCTCTTCTTCTTGGAGCTTCAGACCG
 GGATGAGCCAGGCACAGCCAACTCGGATCTTCGATTCCACATCTGAGCCAGGCTCCAGCCCAGC
 CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGGCTTGGCCCTCAGCCCCAAGGGGAGC
 ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAGGACATGGGTGA
 CCAGGCCTCAGGCCACCAAGGCCATGCCACCGTGAAGTCTCCATCATAGAGAGCACCTGGGTGT
 CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTATACCCGCACCACATGGCCAGGTA
 CACTGGAGTGGGGGTGATGTGCACTATCACCTGAGAGCCATCCCCCGGACCTTTGAAGTGA
 TGCAGAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCAGGCTGAGTACCTGCTCC
 AGGTGCGGGCTCAGAATTCCTATGGCGAGGACTATGCGGGCCCTTGGAGCTGCAGCTGTGGTG
 ATGATGAGAGAATGCAACGTGCCATCTGTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
 CAGTCCACCAGGTACTGAAGTGACTAGACTGTGACGAGAGGATGCAGATGCCCCCGGCTCCCCCA
 ATTTCCACAGTTGTGTATCAGCTCTGAGCCCTGAGCCTGAGGATGGGGTGGAGGGGAGAGCCCTC
 CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTCTCCCATCCGAGAGGCCAGAACAT
 CCTGCTTCTGGTGTGCGCATGGACCTGCGAGGCCAGAGGGTGGCTTCAGCAGCACGTGTGAAG
 TCGAAGTCGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCAGATTGGGCCT
 ATAAGCCTCCCTGAGGATGTGGAGCCCGGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
 CCTCGAGCCCGCTTCCCGCTCATGGATTGTGCCATGAGAGGGGACACAGAAGGACTTTTG
 GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTTGAAGAACCTCAGTTATGAG
 GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCCC
 AGGCCCTGGAGGCCACGCCACGGTGACTGTGCTAGTGGAGAGAGTGATGCCACCCCCCAAGTTGG
 ACCGAGAGAGCTACGAGGCCAGTGTCCCATCAGTGCCCCAGCCGGCTCTTCTGCTGACCAT
 CAGCCCTCCGACCCCACTCAGCCGAACCCCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT
 CTGCATTGAGAAATCTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCCAGCCTGGGG
 ACACCTACACGGTGCCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA
 TACCTCTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCT
 GGCCAGTGGGACCGTTCCTACAGCTTACCCTTGGTCCCAACCCAGGTCAGCAAGGGATTGCG
 GCCTCAGACTCTCAATGGTTCCCATGCCCTACCTTACCTTGGCCCTGCATTGGGTGGAGCCACGT
 GAACACATAATCCCCGTGGTGGTCAGCCACAATGCCCAGATGTGGCAGCTCCTGGTTGAGTGAT
 CGTGTGTGCTGCAACGTGGAGGGGCGATGTCGCAAGGTGGGCGCGATGAAGGGCATGCCCA
 CGAAGCTGTCCGAGTGGGCATCTTGTAGGCCCTGGTAGCAATAGGAATTTCTTCTATCCTC
 ATTTTACCACCATGGACATGTCAAGGAAGAAGGACCCGGATCAACCAAGCAGACAGCGTGGCCCT
 GAAGGCGACTGTCTGAATGGCCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG
 TCCCTTGGGAGAGAGCCAGCACCACCAAGATCCAGCAGGGGACAGGACAGAGTGAAGCCCTCCA
 TCTGCCCTGGGGTGGAGGCCACCATCACCATCACCAGGCATGTCTGCAGAGCTGGACACCACTT
 TATGGACTGCCCATGGGAGTGCTCCAATGTGAGGTGTTTGGCCAATAATAAAGCCCCAGAGAA
 CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
 EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
 YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
 STSLDHALERTYQLLVQVKMDGQASGHQATATVEVSI IESTWVSLEPIHLAENLKVLYPHHMAQ
 VHWSSGDDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL
 VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
 FQVDPTSGSVTLGVLPRLAGQNILLVLAMDLAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
 PISLPEDVEPGTLVAMLTADLEPAFRLMDFAIERGDTEGTFGLDWEPDSGHVRLRLCKNLSY
 EAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT
 IQPSDFISRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTAPVPS
 QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTALHWVEP
 REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTTLVAIGIFLI
 LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCCACGCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
TCTCTTCACCTTCAAGTCCCCCTTTCTCAAGAATCCTCTGTCTTTGCCCTCTAAAGTCTTGGTAC
ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
TGTGACCTCCAGTGGGGTCCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
GCATAGTCACCAACTCTGAGTTCATACAACTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGGCCAGCACAGTCACCAACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCTCCAGTGGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCCACCAATCTGAGTCCAGCACACCTCC
AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGTCCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA
GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCCACCAATCTGAGTCCAGCACAACTCCAGTGGGGCCA
ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGAAGGCAAGCCTGGTGGGTCCCCTGGT
GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTTGCTGGGC
TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT
CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG
GTGGAGTCTTAAGTGGTCTGGAGGAGACAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
ACAGCGGGCCCCTGAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCA
CCCAAGACCTGGTTTCTTTTCAATTCATCCCAGGAGACCCCTCCAGCTTTGTTTGAGATCCTGAA
AATCTTGAAGAAGGTATTCCTCACCTTTCTTGCCTTTACCAGACACTGGAAAGAGAATACTATAT
TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG
CCCCGGGTGGGTATCTAGCTCTGAGATGAATCAGTTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRRPVSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTCCGCACGGAGGTTGTGACCCCTA
 CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGGGCGTGCCCTGCTTGTCACAGGTG
 GGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTCTCTGGGAAGCTGGCAACAAAT
 GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
 AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
 GGTGTCATGGAAAGGATTTACTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGAAGCATT
 TCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC
 CGCCTTGTTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAAAGT
 GATTATAACTGGGGATGCATTTGTTCTGGAGAAAAGAAGTGCATTATCATGAACCATCGGACAA
 GAATGGACTGGATGTTCCCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
 TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGTTGGGCCATGCAGGCTGCTGCCTATAT
 CTTTCATTTCATAGGAAATGGAAGGATGACAAGAGCCATTTGAAGACATGATTGATTACTTTTGTG
 ATATTACGAACCACTTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
 TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
 AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
 ATATCACTGTGGCGTATCCTCACAACATTCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
 CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCCTCCCCACATCCAAGGAGGA
 CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCTTCTATCAAG
 GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCAATCCACCTTGCAAGTCTGAACTCAGGGTC
 CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCCGTTCAGCCCTGCAATGTGCCTACTCAT
 ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
 GAATATTTGGTGGACTGGAGATCATAGAAGTGCATGTTACCGACTTTTACACAAACAGCCACAT
 TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTG
 GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTGCATGACTATGTCGAATATTTCTTACT
 GCCATCATTATTTGTTAAAGATATTTGCACTTAATTTTGTGGGAAAAATATTGCTACAATTTTT
 TTTAATCTCTGAATGTAATTTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
 GGGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTLFWGSFFGSI FMLS P
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCIMRYSYLREKICLKASLKGVPFGFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKSRSNFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFFPREIHFHVHRYPIDTLPTSKEDLQLWCHKRWEEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRVLVVKLLSILYWTLSFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
 TCCAAATCATCCATCCACCCCTGCTGTATCTGTTTTCATAGTGTGAGATCAACCCACAGGAATA
 TCCATGGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT
 CACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCTGCTCCCTCT
 TTCTTGAGACCAGTGCAGAGGCTATGGAAGTGCAGTTTCTCAGGAATCAGTTCCATGCTGTGGTC
 CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAAGTGA
 GTTTGTGAAGGACTCCATTGCAGGGGGCGTGTCTCTCTAAGGCTAAAAACATCACTCCCTCGG
 ACATCGGCCTGTATGGGTGCTGGTTCAGTTCCAGATTACGATGAGGAGGCCACCTGGGAGCTG
 CGGGTGGCAGCACTGGGCTCACTTCCTCTCATTCCATCGTGGGATATGTTGACGGAGGTATCCA
 GTTACTCTGCCTGTCTCAGGCTGGTTCCTCCAGCCACAGCCAAGTGGAAAGGTCCACAAGGAC
 AGGATTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
 ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
 GGTGGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTA
 TTTTACTCGGGTTACTCTGTGGTGCCCTGTGTGGTGTGTGTCATGGGGATGATAATTGTTTTCTTC
 AAATCCAAAGGGAAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
 AGACGCCCGGAAACAGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCGAAGCTCTGCG
 TTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA
 TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
 CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACA
 ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTT
 ACATTCAATCCCCATTTTATCAGCCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCCT
 GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCTGCG
 TGACATGTGAGTTTGAAGGCTTGTGAGACCTATATCCAGCATGCGATGTATGACGAGGAAAAG
 GGGACTCCCATATTATATGTCCAGTGTCTGGGGATGAGACAGAGAAGACCTGCTTAAAGGGC
 CCCACACCACAGACCAGACACAGCAAGGGAGAGTGTCTCCGACAGGTGGCCCCAGCTTCTCT
 CCGGAGCCTGCGCACAGAGAGTCAACCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC
 TGAGCCCTGCAGCAGCGGCAGTCCAGCTTCCAGATGAGGGGGATTGGCCTGACCCTGTGGGAG
 TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA
 CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTGCTAGTCACGG
 ACAGTGATTCTGCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT
 TGAGGGCACAGTGTGCTAATGATGTGTTTTTATATTATACATTTCCACCATAAACTCTGTT
 TGCTTATTCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACACC
 TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTCACTGATTCTATAAGCCCAGCAT
 TACCTGATACAAAACAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC
 CTCATTAACACAGACACAAAATTTCTAAATAAAATTTTAAACAAATTAACCTAAACAATATATTTA
 AAGATGATATATACTACTCAGTGTGGTTTGTCCACAAATGCAGAGTTGGTTTAAATATTTAAAT
 ATCAACCAGTGTAATTCAGCACATTAATAAAGTAAAAAGAAAACCATAAAAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
 LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
 VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGDLSSDSRANADGYSLYDVEISI
 IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK
 SKGKIQAELDWRKKGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
 TRKSVVASQGFQAGRHYWEVDVGQNVGWYGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
 FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
 TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCAAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
 TTTGTTGGGAACCCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
 CTCCTGTTTATTATGTGAGATATAATCAAAAGAAGACCTACAATTACTATAGCACATTGTCATTTACAAC
 TGACAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAAATTTTACAGAAATGAGCCAGAGACTTGAAT
 CAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
 AGTCAACAGAAGCATGGAGTGTGGGTCATATGCTGTTGATTGTAGATTTCTACTCTACTGAGGATCCTGA
 AACTGTAGATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCCTAAAGTAG
 ATCCTCACTCAGTTAAAATTAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA
 ACACGAAGAAGTAAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG
 GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCAATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC
 TTGTGAGTGTCTCACTGTTTTACAACATATAAGAACCCTGCCAGATGGAGTGTCTCTTTGGAGTAACA
 ATAAACCTTCGAAAATGAAACGGGTCTCCGGAGAATAATGTCCATGAAAAATACAAACACCCATCACA
 TGACTATGATATTTCTCTTGACAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTG
 TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAAT
 GATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
 TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
 GCCAGGGTGACTCTGGAGGACCCTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG
 AGCTGGGGAGATGAATGTGCGAAACCAACAAGCCTGGTGTTTTATACTAGAGTTACGGCCTTGCGGGAGTG
 GATTACTTCAAAACTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTTG
 GGTGTGGAGGCCATTTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTGACTGATCTCA
 ATAACTGTTTGCTTGATGCATGTATTTCTTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA
 GATCAACTCTGTCTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATAATACAAATATTAC
 ATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTTGTGCAAGTTTTGACTTGTGACATAAATTTGTAAT
 GCATATATACAATTTGAAGCACTCCTTTTCTCAGTTCCTCAGCTCCTCTCATTTCAGCAAATATCCATTT
 TCAAGGTGCAGAACAGGAGTGAAAGAAAATATAAGAAGAAAAAATCCCCACATTTTATTGGCACAGAA
 AAGTATTAGGTGTTTTTCTTAGTGGAATATTAGAAATGATCATATTCAATTATGAAAGGTCAAGCAAAGACA
 GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAAGTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
 ATATATCCTTATTTTCATTTCCAAACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTGTGACCT
 ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATATTATTTAACATTG
 TTAAGTGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTNYYSTLSFTTDKLY
 AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
 PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
 TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
 RRIIVHEKYKHPSHDYDISLAELSSFPYTNVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
 QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDIWYLAG
 IVSWGDECAKPNKPGVYTRVTALRDWITSKTI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCCTG
 CCCCAGATGAGCCCCCGCGTCCCGGACTATCCCCAGCGGGCGTGGGGCACCAGGGCCAGC
 GCGGACGATCGCTGCCGTTTTGCCCCCTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT
 ACGGGGCTCACAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA
 TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGGACTACCTAA
 ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTGACTTACTTTCTT
 GTGGTTTCATCCGGTCATGATTGCTGTTTCTGCTTCTTATCATTGTGGGGATGTTAGGATATTG
 TGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCATGGTACTTTGGAAGTTTCTGCTGCTATTTCT
 GTGTAGAAGTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTATGGTTCCAGTACAATGGTCA
 GATATGGTCACTTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA
 TGCTTGGAATTTTTTTCAGAGAGAGTTTAAAGTCTGTGGAGTAGTATATTTCACTGACTGGTTGG
 AAATGACAGAGATGGACTGGCCCCAGATTCTGCTGTGTGTAGAGAATTCAGGATGTTCCAAA
 CAGGCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTTGTGGGAAGAAAATGTATTCCTT
 TTTGAGAGGAACCAACAACCTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA
 TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
 ACAGACCAATGATGTCCTTGAAGAATGACAACCTCAGCACCTGTCTGTCCTCAGTAGAAGT
 GTTGAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT
 TTGAGATGGAGGAGTTATAAAAAGAAATGTCACAGAAGAAAACCAAACTTGTTTTATTGGACT
 TGTGAATTTTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAAAATGTTGCCATAAAA
 TAACACCTAAGCATATACTATTCTATGCTTTAAATGAGGATGGAAAAGTTTCATGTCATAAGTC
 ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
 TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCCGCA
 TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGTTGATTTACTTCTACCAA
 CTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
 CTCAGCGATCTATTCTCTGATGCTAAATAAATTATATATCAGAAAACCTTCAATATTGGTGACT
 ACCTAAATGTGATTTTGTCTGGTTACTAAAATATTCTTACCCTTAAAAGAGCAAGCTAACCAT
 TGTCTTAAAGCTGATCAGGGATTTTTGTATATAAGTCTGTGTAAATCTGTATAATTCAGTCGAT
 TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATT
 ATTTTTAGCCTTTCCGTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC
 TGTTATTTAAATACTTAACCCTAATTTTGAATAATACCAGTGTGATACATAGGAATCATTATTC
 AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTACAGA
 AAGGACTTGTATGCTGTTTTCTCCCAAATGAAGACTCTTTTGGACACTAAACACTTTTTAAAAA
 GCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTTACAGAAAA
 TAGTGTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAACATGTGACATTTAGAGATT
 CTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAATTTTTTACAA
 GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTATTTTGTATTAT
 TTCTCAGAAATATGAAAAGAAAATTAATGTGTCAATAAATATTTCTAGAGAGTAA

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTCTCTCTGATAAAGCCCTACCAGTGCT
GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAGTGCTTGAAAGAGAAGGGGACAAAGGAACA
CCAGTATTAAGAGGATTTTCCAGTGTCTTGGCAGTTGGTCCAGAAGGATGCCTCCATTCTGCTTCTCACCTG
CCTCTTCATCACAGGCACCTCCGTGTACCCGTGGCCCTAGATCCTTGTTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACAGTTGGATGAGTCTCAAGGTCTCTCTATGTGACAACCATGTGAATGGGGAG
TGGTACCACTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAAGGCGACGGCATGTGCAACGCCAGGCTTGTGCCAGCT
TCAATGGGAACGTGCTGTCTCTGGAACACCAAGGTGGAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGAGTGCACATGCGCTCCAGGAAGTGTGCTAGGCCCTGACAGGCAGACATGCTTGG
ATGAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCTACCGCTGT
GAGTGTGGGGTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTTGCCCTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCCGGGGCCTGGTGTGT
CTGAGGATAACCACACTTGCCAAGTCCCTGTGTTGTGCAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGGTTGGTGGCCTGGAGCTCTTCTGACCAACACCTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTTCTCTCAAGACATGTGGTACAGTGGTGCATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
CAGGTCTACCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG
ACCTGCGAGTTTCCACGCCTGTACACCATTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGGATCTTCCCATTCACCTCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC
GGGAAGCTCTGCCACCTCAAGCTTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
TTGGAAAGCTTGGTGGAGAGCTGCTTTGCCACCCCACTCCAAGATCGACGAGGTCCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTTCAAGTGTGGGCAAGACCACAAGGAAGTGTCTTGCATGCGCGGTTCTTGTCTGTGGAGTG
TTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGGCGAATGCGTCTGGGGCAGGAGGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGGCGGCGGATCCGCATCGACTGGGAGGACTAGTTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACCAGCCCTCTAAGAACATCTGCCAACAGC
TGGGTTTCACTTTCACACTGTGAGTTCAGACTCCCAGCACCAACTCACTCTGATTCTGGTCCATTCACTGGGCA
CAGGTTCACAGCACTGCTGAACAATGTGGCCTGGGTGGGGTTTCATCTTTCTAGGGTTGAAAATAAATGTCCA
CCCAGAAAGACACTCACCCATTCCCTCATTTCTTTCCTACACTTAAATACCTCGTGTATGGTGCAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCTAGAAAAATTAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAATTTCAATTCAATGCAGACTAA
TTATAGGGAATTTGGAAGTGTATCAATAAACAGTATATAATTTT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYVYRLTKPSVCFHV
YCGHFDICDEDCGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDEVEGCHNNNGGCSHSLGSEKGYQCECPGRLVLSNHTCQVPVLCKSNAIEVNIPRELVGG
LELELTNTSCRGVSNCTHVNILFSLKTCGTVVVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEPEPYREALPTLKLRLDSLYFGIEPVVHV
SGLESLVESCFATPTSKIDEVLKYYLIRDGCVSDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRGAGGEDSAGLQGQTLTGGPRIIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
CCTGCTGCCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
GGTTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCCAAGGGCCTGCAGTGGCAGTCCGCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAAGTGGTTCTCTGCTGTTTCGACAATTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGCATGCGGAATCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCTG
TGGGAAGAGCCTGAAGACCCCCGTGTGGTGGGTGGGGAGGAGGCTCTGTGGATTCTTGGCCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA
GCCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAG
CTTCCCATCCCTGGCTGTGGCCAAGATCATCATTTGAATTCAACCCATGTACCCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTTCCTCACTCACTTTCTCAGGCACAGTCAAGGCCATCTGTCTGCCCTTCTTT
GATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATTTGGATGGGGCTTTACGAAGCAGAATGGAGGGAA
GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCGT
ACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGAAGGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCCGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
CTGCGGGGGCCCCGAGCACCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCTGCCCTTTGCAAGTCTGGGAGCCGCTTCCTTCTGCTGCCCTGCCACCT
GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCAT
TTCTTGGAGCAGCAAAGGGCCTCAATTCCTGTAAGAGACCCTCGCAGCCAGAGGCGCCAGAGGAAGTCA
GCAGCCCTAGCTCGGCCACACTTGGTGTCTCCAGCATCCAGGGAGAGACACAGCCACTGAACAAGGTCT
CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAAAGCC
CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA
GCCTACTAGAGCAAGAAACAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTCATTGTTATTACAGCTATGGCCACTATTATTAAGAGCTGTGTAACATCTCTGGCAAAAAAA
AAAA

FIGURE 112

MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVLIKVILDKYYFLCG
 QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
 FTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
 KTPRVVGEEASVDSWPWQVSIQYDKQHVCSSILDPHWLTAHCFRKHTDVFNWKVRAGSDKL
 GSFPSLAVAKIIIIIEFNPMYPKDNNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWG
 FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCQGDSGGPLMYQS
 DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTCAGCAACTAAAAAAGCCACAGGAGT
TGAAGTGTAGGATTCTGACTATGCTGTGGTGGCTAGTGTCTCTACTCTACCTACATTAAAAATC
TGTTTTTTTGTCTCTTGTAACTAGCCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCTCGAAG
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGCTAGCTGCGGCTTTCAAGGTGGGC
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCGAAGATTTCATAGGCGATGGCTCCCACTGCCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCTGGGGCCAGGACGGGCGGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGTGCCCCCATCTAACCTTTTCATGTCTGCACATCACCTG
ATCCATGGGCTAATCTGAAGTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCATTCTCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACCTTGTGCCAAATTATGGGTGAGAAAAGATG
GAGGTGTTGGGTATCACAAGGCATCGAGTCTCTGCATTCACTGGACATGTGGGGGAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTTGGGGCCATCAGACAGCCGTTTCCGCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTCTCTCTCC
CTCCTTCCCTCTGAGAGGCCCTCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCCAGGAGGTGAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTGAATTTGAACCCCAATCCA
AACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT
TTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCTGGCCAACATGG
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTA
TGGTTATTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:
amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
 AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
 ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCT
 GTTTTGGGGGAGCAAGCACTTCTGGCCGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
 TCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAAGTGAATA
 TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
 CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCTGAATTTT
 CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACTTCTTTGAACAGTCAGTG
 ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCCAAAATCTGGA
 GATTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG
 ACTTTGAGGAGGAGGAGAAGATCTTCACTTTCCTGCCAACGAAAAAAGGGATTGAACAAAAT
 GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA
 AGAACTTCCAATAAATGACTATACTGAAATGGAATAGAATTGATCCCATGCTGGATGAGAGAG
 GTTATTGTTGTATTIACCTGCGGTCGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTA
 GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG
 TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAAC
 TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT
 GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAATTCTACA
 TTATCACCAAAAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWPVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA
GGGCGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCCTCCTGTCCATCCTGGGGCT
GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTCACT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCC
TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGTCTT
GGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG
TCTGTGTTTGCCAAACATGCTGGTGACTAATCTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
GATGGTGCAGACTGTTTCCAGACCAGGTACACATTTGGTGCGGCTCTGTTCGTGGGCTGGGTGCTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGCGCTGGCACCAGAAGAAACCACTACAA
GCCGTTTCTTATCATGCCTCAGGCCACAGTGTGCCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCACGACTATGTGTAATGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA
CCCCAAAAACAAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
CGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCACCATAAAACA
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC
TTCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACTTTGAAGGAAA
GAGTAGACCCAAAGATGTTATTTCTGCTGTTTGAATTTGTCTCCCCACCCCAACTTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT
ACACTGTGATCTTAAAAGTTACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAAACCTTCTACTGCTG
TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTTCTGT
CGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAATTTAAGTCCTAAATATAGTTAA
AATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATCC
ATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGT
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAATCAGCCAGTCA
TGGTGGCATAACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTAAAATAATCTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
 YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
 CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
 CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
 V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCGGAGTCC
 AGCTGGCTAAAACATCATCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTTCTTG
 GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACGTGCCTCAGTGGAGAGTGTGGCCTTCATT
 GAAAAACAACATCGTGGTTTTTGGAACTTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
 CATCAGGATGCAGTGC AAAATCTATGATTCCTTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
 GACTGATGTGTGCTGCTTCCGTGATGTCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC
 ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
 CATCACGGGCATGGTGGTGCTCATCCCTGTGAGCTGGGTGCCATGCCATCATCAGAGATTTCTATA
 ACTCAATAGTGAATGTTGCCCAAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA
 CTGGTGCTGATGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTGTGCAACGAAAAGAGCAGTAGCTA
 CAGATACTCGATACCTTCCCATCGCACAACCCAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG
 TCTACTCCAGAAGTCAGTATGTGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATG
 ACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAAGAACTTTGATTTACTGTTCTTAACTGCCT
 AATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACAGCAGAATGAGATA
 TTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
 CTCTTTTTATCATTTACTTCAAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
 ACGACATAGCATTATGTACATAGATGAGTGTAACTTTATATCTCACATAGAGACATGCTTATATGGT
 TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA
 ATCATGGATAGGGTTGAAGAAGGTTACTATTAATTGTTTTAAAAACAGCTTAGGGATTAATGTCCTCCA
 TTTATAATGAAGATTAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT
 GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTT
 CTTGTGTATTAAATTAACATTTTTTAAACGCAGATATTTTGTCAAGGGGCTTTGCATTCAAAGTCTT
 TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTG
 AAAATATTTTTGTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGAT
 ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATAATAATAAAGAGCAGAAAAATA
 TGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAACTTCACCT
 GCTCCTATGTGGGTACCTGAGTCAAAATTGTCATTTTTGTTCTGTGAAAAATAAATTTCTTCTTGTA
 CCATTTCTGTTTAGTTTTACTAAAATCTGTAAATACTGTATTTTTCTGTTATTCCAAATTTGATGAA
 ACTGACAATCCAATTTGAAAGTTTGTGTGACGCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT
 TATACATTTATATTAATAAATTTGTACATTTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQCK
 IYDSLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
 MVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCNEKSSSY
 RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGGAG
CCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCGGCAGCC
GGGAGCCATGCAGACCCAGGGCCCCGCGCCTCCCCGCAGCGGCTCCGCGGCCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGT'CGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCCGGGATG
GATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGC'TTGAGGAGTCCTGGACACCCAAC'TAC
AAGCAGTGTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGTACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTCAGTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTCAAGACCTCTTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGTGGATTAGTGGATGTTGCTATCTGGG
TTGGCACTTGTTAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAACTACCAAAATTAAATGCTTTAATTTTCATTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTG
CTTCAATCAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGT'TTTTCTCTTAGTATAGCATTTTTA
AAAAATATAAAAGCTACCAATCTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
 DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
 MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGSLPIEAIYLDQGSPEMNSTINIHRTS
 SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
 CTGGGTGCTCATCACGGGAACCTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
 ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTA
 ACCGCCCCCTCCCCACCCCCCAAAAAAAGCTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC
 TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG
 TGTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCAAGGGGTCCAATTTT
 TCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCATGCAACTG
 GCCCCAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGTTTCAATG
 TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCCACTGTCTTACTGACAATG
 CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA
 ATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
 ATAACAGCCTTCAAAAACCTTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC
 CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTTAATGGAATACGCAGACTCAAAGA
 GCTGATTCTTAGTTCCAATAGAATCTCCTATTTCTTAAACAATACCTTCAGACCTGTGACAAATT
 TACGGAACTTGGATCTGTCTTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG
 CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA
 AGACTGCCGCAACCTGGAACCTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG
 TCTTTGCTGGCATGATCAGACTCAAAGAACCTTCACCTGGAGCACAATCAATTTTCCAAGCTCAAC
 CTGGCCCTTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGAATAAAATCAGTGT
 CATAGGACAGACCATGTCTGGACCTGGAGCTCCTTACAAAGGCTTGATTATCAGGCAATGAGA
 TCGAAGCTTTTCACTGGACCCAGTGTTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGAT
 TCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGATTCTTGATATCCCTCAATGACATCAG
 TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAACCTGGCTGAAAAGTT
 TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
 ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAG
 GGCTCTCCCAAAGCCGACGTTTAAGCCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT
 TGCCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
 TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCCTGCTGGTTAT
 CTACGTGTCATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
 GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTTTATGTA
 GATTATAAACCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTA
 TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTGTGATAAAAAGAGCTCTTAAAGCT
 GGGAAATAAGTGGTGCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCTCCCC
 TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCTTGTCCGTTTTAGTGCATTACATAACT
 GGTCAATTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
 GAACCTCCGGTTAATATAATACCTATTGTATAAGACCCCTTTACTGATTCCATTAAATGTGCAATTT
 GTTTTAAGATAAACTTCTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTLLTMLSSAERGCPKGCRCGKVMVYCESQKLQEIPSSISAGCLG
LSLRYNLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKLHLEHNQFSKLNLFRLVSLQNLQWNLKISVIGQTMSTWSSQLRLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIGQEILDSWISLNDISLAGNIWECSRNICSLVN
WLKSEFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMQLQQR
SLMRRHRKKRQSLKQMTPTSTQEFYVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGAATGTCGTCCTCCGAGGAGGAGGAGAGGCTTTTGCCGCTG
 ACCCAGAGATGGCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCGAGCT
 AGCAACCTTTCCCCTGGATCTCACAAAACCTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC
 GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC
 ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT
 GTATTCTGGAGGTGCAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG
 ATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT
 TTAGCCAATCCAACCTGACCTAGTGAAGGTTGAGATGCAAAATGGAAGGAAAAGGAACTGGAAGG
 AAAACCATTGCGATTTTCGTGGTGATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC
 GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA
 ACCACTTATGATACAGTGAACACTACTTGGTATTGAATACCACTTGAGGACAATATCATGAC
 TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
 TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG
 ACTGACTGCTTGATTCAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTACC
 ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTTCTGGCTTACTTATGAAAAATCAGAGAGA
 TGAGTGGAGTCAGTCCATTTTAA

FIGURE 126

MSVP EEEERLLPLTQRWPRASKFLLSGCAATVAEELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGII EEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVF GKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLE GKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYD TVKHVYLVNLTPLEDNIMTHGLSSLSGLVASILGTPADV IKSRI MNQP
RDKQGRGLLYKSSDCLIQAVQGEGFMSLYKGFLPSWLRMT PWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 128

MARPGMERWRDRLALVTGASGGIGA AAVARALVQQGLKVVG CARTVGNIEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPD TLLSGSTSGWKDMFNVNVLALSICTR
EAYQSMKERNVDDGHIININSMSGHRVLP LSVTHFY SATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACCTCTACATGGGCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
 CATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTCAAGTGCAGAGTCTCAGTTGCCCGGGAGC
 ACCTCCCTCCCGAGGCAGTCTGCTCAGAGGGCTCGGCCCAGAATCCAGTTCTGGTTTCATGC
 CAGCCTGTAAAAGGCCATGGAACTTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA
 GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAGCCACCTGTGG
 CCGTTGAGTGCTTGAAATGAGGAACTGAGAAAATTAATTTCTCATGTATTTTTCTCATTTATTTA
 TTAATTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA
 TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTATTTCTTT
 TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
 CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
 GCACCACAATGCCCCAACTAATTTTTGTATTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
 CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTTATGATTACA
 GCGGTGAGCCACCGTGCTGGCCTAAACATTTATCTTTTCTTTGTGTTGGGAACTTTGAAATAT
 ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACTGGGACTTCTTCCCTCT
 ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC
 AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCATGTG
 AGTAAGAAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG
 TTCCATCCATGTTGCTGCAAATGACAGGATTTGTTCTTAATTTCAATTAAAAATAACCACACATG
 GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KKGHTLGESPMFVKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTACTGTAATGATCGCTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACCTCTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCCTTCAGAT
TTGAAAACTTGCTGAAAGTAGAAAGAATATACCTTATACCAACAACAGTTTAGATGAATTCCTACCAACCT
CCCAAGTATGTAAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
AAATTCCCTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCCGACTGCTTTTCCGTCCCGTAATCACCTTAGCACAAATCCCTGGGGTTT
GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTATCACCATCTCTTCAAG
GTCTCACTAGTCTAAAACGCCCTGGTTCTAGATGGAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT
TTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATTCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCAAATGCTTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT
GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCTGTGGGATGGCTATTAAAGGATCTCAATGCAGAACTGTTTGATGTGAAGGACAGTGGGATTGTAAGC
ACCATTCAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCCCAAGGACAGTGCCAGCTCCAGTGAC
CAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAACCAAGGGAGTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTATATCTCTTGGAACCTTGCTCTACCTATG
ACTGCTTTGAGACTCAGCTGGCTTAAACTGGGCCATAGCCCGGCATTTGGATCTATAACAGAAACAATGT
AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTCACCCTATAAAGTATGCATGGTTT
CCATGGAAACCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTT
CGAATGTACAACCCTACAACCACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACC
TTTGGCTGCCATCATTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTGGTATG
TTCATAGGAATGGATCGCTCTTCTCAAGGAAGTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAAATCAGGGAAACTTCTTTTCAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAATACACACCATATTTCTCCTAATGGAATGAATCTGT
ACAAAAACAATCACAGTGAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTGGGTTTTTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCNDRFLTSIPTGIPEDATTLYL
 QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
 EELHLDNDSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
 QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
 RVPPNAFSYLRQLYRLDMSNNNLSNLPQGI FDDL DNITQLILRNNPWYCGCKMKWVRDWLQSLPV
 KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
 IKNPKLTKDQQTGSPSRKTTITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
 IVTGERSEYLVTALEPDSPYKVCMPMETSNLYLFDETPVC IETETAPLRMYNPPTTLNREQEKE
 PYKNPNPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
 ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLKNNHSESSSNRSYRDSGIPDSHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
 640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 134

MSARGRWEGGRRACRGSGLGLARAQGAERTVSSEQRPAMASLGLLLLLLLLTALPPLWSSSLPGLD
 TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLOPLSLRV
 GMLGEKLEAAIQRSLSHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPDQSFSEE
 RSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
 YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSQKQEGCFG
 EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSLILP
 SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTC'TGAGTGCAGAGCTGCTGTCAATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCCGAGTGTGTTAGACGGAGAAGAGCACGTCCGGTTTCCTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTTAAAAGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAATGGT
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCGAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAATCACTCTGGCAAATCTAGCAGCGGCAGCAG
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTTCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACGTGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT
GAGGTCAATATTGATGTCACCTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT
ATGAACTACTATACATTATGTATATTAATTAAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEGSGSGVGIGDRFKIEGRAVVPGVKPQDWISAA
 RVLVDGEEHVGLKTDGSFVVDIPSGSYVVEVVSPAYRFDPVRVDITSKGKMRARYVNYIKTSE
 VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLI FVLLPKVVNTSDPDMRREME
 QSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGGGA
 CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCGCCAAAACAA
 GTTTTGACATTTCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
 CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
 CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAGGCTTGGCAGT
 TTTTCTTACTCTGTGGTCTCCAGATTTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCCTTCAGC
 CTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
 AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
 TGCAAGCCAAAGATGGAAACATTGACATCAGAACTTAAGGAGGACTGAGTCTTTGCAAGACACA
 AAGCCTGCGAATCGATGCTGCTCCTGCGCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAA
 AAATACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA
 CCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
 AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
 TTTGGGGGAACTAGACATTCTTCTGCAATGGATGGAGGAGACAGAAATAGGAGGAAAGTGATGCTG
 CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
 ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
 CTTCTTGCATGATTGTCTTTATGCATCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
 TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTGTATTTA
 ATGTATTTATTTTTTACTTGGACATGAACTTTAAAAAAATTCACAGATTATATTTATAACCTG
 ACTAGAGCAGGTGATGTATTTTATACAGTAAAAAAAACCTTGTAATTCTAGAAGAGTGG
 CTAGGGGGGTATTTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
 ATTTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTTTGATGTGGAATTGCAC
 ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAAATTGTGTATCTTCCAG
 CCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTCTTTGCATACCAAAAAAA
 AAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPG LTCWALTAEPGWGQNGATT CATNSHSDSELRPEIF
 SSREAWQFFLLWSPDFRPMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
 FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
 SLANSFLT IKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
 E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCCGCATCCAGCC
TAGCGTGTCCACGATCGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCCTTCGGGGATTCTTCCGGCTCCCGTTCGTTCTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCAGCGCCGAACCTTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC
ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCCTGAGAGATGATTTTGTGTTG
GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTTACCTTGTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCCTGGCTTTGTCGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGGAAGACAGTGTGA
TAAGACAAGCAAAAGCAGCTGGAAAAAGAATAGTCTTTTATGGAGATGAAACCTGGGTAAATTA
TTCCCAAAGCATTGTGGAATATGATGGAACAACCTCATTTCGTGTCAGATTACACAGAGGT
GGATAATAATGTACGAGGCATTGGATAAAGTATTAAGAGAGAGATTGGGACATATTAATCC
TCCACTACCTGGGGCTGGACCACATTGGCCACATTCAGGGCCCAACAGCCCCCTGATTGGCGAG
AAGCTGAGCGAGATGACAGCGTCTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA
GACGCCCTTACCCAAATTTGCTGGTTCTTTGTTGGTGACCATGGCATGTCTGAAACAGGAAGTCAG
GGGCTCCTCCACCGAGGAGGTGAATACACCTCTGATTTTAATCAGTTCTGCGTTTGAAAGGAAA
CCCGGTGATATCCGACATCCAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCCTATTCCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAACTGTTGCAAGAG
AATGTGCCGTATATGAAAAAGATCCTGGGTTTGAGCAGTTTAAAATGTCAGAAAGATTGCATGG
GAACGGATCAGACTGTACTTGGAGGAAAAGCATTGAGAAGTCTATTCAACCTGGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAG
TTCTCACCCCTGCTCCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCA
CTGTCACTCCTGGGTTTCTCTGCTCTTTTATTGGTGATCCTGGTTCTTTTCGGCCGTTACGT
CATTTGTGTCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCTCTCGTGGCTGGCGGCAGGCT
GCCTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGTCAGTGGCCTGGAC
AGGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCCAGGCCTCTGGGTGTCCCACACAGGTG
TTCACATCTGTGCTGTCAGGTGATGCCTCAGTTCTTGGAAAGCTAGGTTCTGCGACTGTTAC
CAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA
TCGGACAGCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTTCGCACCACTCATCTGCCACCCCCAGAATGCATCCT
GCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTTCTGTTGGAATCTTAGTCCTTGCCC
TCGGACACCTTCATTTCGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC
ACCCCAACCCTGCACAGCCCTCATCCCCTCTTGGCTTGAGCCGTGAGAGGCCCTGTGCTGAGTGT
CTGACCAGACACTCACAGCTTTGTCATCAGGGCACAGGCTTCCTCGGAGCCAGGATGATCTGTG
CCACGCTTGCACCTCGGGCCCATCTGGGCTCATGCTCTCTCTCCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDVFVFGSKGVKEMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDHHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTACTGCTGCTGTTGGAGTACAAC
TCCCTATAGAAAACAACAGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATTCAT
AGCAGTTCCAGATAAAAACTACATACGCCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCCAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCAACCAGTTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAACTGCCCCATTGAACGCCTTCCTCGCTA
ATTTGAACTAATTGTATAAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLD SGNLI AVPDKNYIRPEI
FFALASSLSSASAEKGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAAQKESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

CTGAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTTTAA
TCCAGGATCCTGTCTTCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAGTTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGGG
CACGGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCACAGTCAGCACCATGCTGGGGCTGCCGTGG
AAGGGAGGTCTGTCTTGGGCGCTGCTGCTGCTTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTTCACGAGCAAAGGGACTGTGATGAACACAATGTATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCACAGACTGGGG
CACATCTTGAATTCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGGAAATTTGAAGACGACATTGACAACCTGCCATTTCCAAGAAAGCACAG
AGCTGAACAATACTTTACCTGCTTCTTACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTGAGTGAAACCCACTCACAGGCTTGTCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT
TTGGACTTGTTTTGTTATCCTATTTTGCATGTGTTTGAGATCTCAGATCAGTGTTTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAACATCAGCATTTTAAGAAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCCAGGTGGCTGAGAACCGCCCGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTCACCGGCTGCATCAATGCCA
CCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTCACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCCAAACCCACGCGTGTCTGAAGGTGCCCAGGAGCGGCGATGCACTCGCACTGCAAA
TGCCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTCTTAGAGGTAG
GTGTATTCCCGTTTACAGATAAGGAAACTGAGGCCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCTGGCTTGTCTAACCCAGGTTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
 LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ
 VLWRLVQELCSLKHCEFWLERGAGLRVTMHQPVLLCLLALIWLMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATTT
CCTGATGATTTATAGACTCAAAGAAAACTATGTTCAGAAGCTCTCTTCTCTTCTGGCCTCCTCT
CTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTTCATTTTGCTTGTCAGTGGGGTAGGTCAGTCTTAGTTTTATTTTTGAAATTT
CAACTTTCAGATTCAGGGGGTACATGTGAAGGTTGTTTTATGAGTATATTGCATTGATGTCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLILISSIYSECKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCCGCGCGCCCCAACCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCTGCAGCCGCCCAGTCC
CGGCCCCCTCTCCCGCCCCACACCCACCTCCTGGCTCTTCTGTCTTTTACTCCTCCTTTTCATTCATA
ACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGGCTGTGACCCAAGCCGAGCGTGAAGAATGGGGTT
CCTCGGGACCGGCACTTGATTCTGGTGTAGTGTCTCCGATTCAAGCTTTCCCCAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAAITTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTGAGAGCAACTATCTTTT
TGTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACTG
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGGACTCCTTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTATGAAG
AAAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAAACTACTTAATCTCGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAAATTAATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCCAATAAGCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGACAACCTTTGAGGAACCTCA
ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAAGAACAAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCT
CCAGAAGAAGGTGTTTCTTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAAACAAGCT
AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC
CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAGGGAATAAAGAAGATTATGACCTTT
CAAAGATGAGAGACTTCATCAATAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAATGGCAAAAGATCCAGGAGTCTTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATTCTGTGATTAAAAATTTTGGACCAAGG
GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAAACATAGCTTCTTCCC
GTAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPKVTMAAIQDGLAKGENDETVSNTLTLTNGLE
RRTKTYSEDNFEELQYFPNFIYALLKSIDSEKEAKEKETLITIMKTLDIVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLKMRDFINKQADAYVEKGILDK
EEAEAIAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESISVVPNRWLDASLSPVILGVQGGG
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENG GW NAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTCTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTTCACCCTTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTT
CAGCCTTATATGCAGGAGGTGGTGCCCTTCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTTGCTGTTTATGTCTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCCTGCTAGAAATAA
CAATTAGATGCCCCAAAGCGATTTTTTTTAACCAAAAGGAAGATGGGAAGCCAAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAACCAATGCCACTTTTGTTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTTATTGTAACCTGGTGTTT
TATACACAGAAAACAATTTATTTTTTAAATAATTGTCTTTTTCCATAAAAAAGATTACTTTCCAT
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTCCATAATCAGTACTTTATATTTATAAA
TGTATTTATTATTATTATAAGACTGCATTTTATTTATATCATTTTATTAATATGGATTTATTTAT
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLLVQGGAAPISSHCRDKNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFHGVSMSERCYLMKQVLNFTLEEVLPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GCTTGTCTGAAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTCAGTGCAGAGGGC
TGCTTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGG
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCTACAGGTGGTTGCAT
TCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG
GCCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACCACGCCCGT
TGCTGTGCCCGCACTGCGTCAGCCTACAGACAGGCTCCCACATGGACCCCGGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCGAGCGCAGGCTGTACCGTGTTTTCTTAGCTTGTGTGTGTGCGGCCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCCGTGAAGTGCTGTCTGGAGCAG
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAAAACCTGCACTTCTGCACATTTTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGGAAGCTGGTGTCTGTCAATTTTCTCTCAGGAAAGGTTTTCAAA
GTTCTGCCCATTTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTTCCCATCCCCTGCTACCCTG
GCCAGCACAGGCACTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCCTGTTTATTAT
TGTTTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPSKGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHTKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
 GACCGTTCAATGTGGCTCTGAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
 CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACAGTGTGCAACAGGGGACTATTCA
 ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGAAGGCCACCAA
 GATTTGTGTGACGGGCAAAAGCAACTTCCAGTCTACAGCTGTGTGAGGTGCAATTACACAGAGG
 CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCTACATCGGCTTCCCTGTA
 GAGCTGAACACAGTCTATTTTATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG
 CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAA
 AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGAAGAAGATGAGGAGACA
 GTAGAAGTGAAC TTCACAACCACTCCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
 TATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
 TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCTACTTGTGGC
 AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCT
 GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTCTGCTGGTGGCCA
 CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
 TCTACCACCACACTACTGCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA
 TCACACAATTTGTTACTTCACTGAATTTCTTCAAACCATTCAGAGTGAGGTGATCCTTGAAA
 AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCA
 GCAGACAAAGTCGTCTTCTTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
 GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCA
 GTGATCTAAGAAGCCAGATTATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
 AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
 CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
 ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGFSPEWMLQHDLPGLRDLRVEPVTTSVATGDYSILMNVSWV
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFYSYIGFPELNTVYFIGAHNIP
NANMNEDGPSMSVNF TSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK
SKPGGWLPLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKIAEMGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGPSSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKKQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
 GTCGATATTGGGGCTTGCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
 CTTTTTCCAAAAGCCTGAGAGTTGCCCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
 ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
 GAATTACACTGTCACCTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
 ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA
 GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTCCAGTTGGAGAAGGTGCT
 GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTTAAGAGGTGCATATCC
 ACTCAGCTGAAGAAG

FIGURE 16o

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVSMsrNIESRSTSPWNYTWTWDPNRYpSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCsvSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAG
ACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGCTGAAGAGGGATTC
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGCTGCCCCCCCTGGGGGGGGCAGCAC
AGGGCTCAGGCTGGGTGGCCACTGGCACCTAGAAGATGCTGCCCTGGTTCTTGCTGTCTCT
TGGCACTGGGCGGAAGCCAGTGTCTCTTCTCTGGAGAGGCTTGTGGGGCCTCAGGACGCTACC
CACTGCTCTCCGGGCTCTCTGCGCCCTCTGGGACAGTGACATACTCTGCCTGCCTGGGGACAT
CGTGCTGCTCCGGGCCCCGTGCTGGCGCTACGCACCTGCAGACAGAGCTGGTGCTGAGGTGCC
AGAAGGAGACCGACTGTGACCTTGCTGCGTGTGGCTGTCCACTTGGCCGTCATGGGCCTGG
GAAGAGCTGAAGTAGGAAAAAGTTTGGAGGAGCAGCTGACTACGGGGTGGAGGCGCTAGGAA
TGCTCTCTCCAGGCCCAAGTCGTGCTCTCTTCCAGGCCTACCCCTACTGCCCGCTGCGTCTTG
TGGAGGTGCAAGTGCTGCTGCCCTTGTGCAGTTTGGTCAGTCTGTGGGCTCTGTGGTATATGAC
TGCTTCCAGGCTGCCCTAGGGAGTGTAGGTACGAATCTGGTCTTACTACTGACCCAGGTACGAGAA
GGAACTCAACCACACACAGAGTGCCTGCCCTGCCCTGGCTCAACCTGTACGAGATGGTGACA
ACGTGCATCTGTTCTGAATGTCTCTGAGGAGCAGCACTTCGCCCTCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCCAAACCCCGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTTGAA
CCACACAGACCTGTTCCCTGCCCTCTGTATTAGTGTGGCCCTCTGGAACCTGACTCCGTTAGGA
CGAACATCTGCCCTTCCAGGAGGAGCCCTCCGCGCACACAGAACCTCTGGCCAGCCCGCAGTGC
GCTGCTGACCTGCAGAGTGGTCTGTGAGCACCCTGCTGCTGTGGCAGAGAGCGGCATCT
GTGCTGGCGGGCTCCGGGTGGGGACCCCTGCCAGCCACTGGTCCCAACCGCTTTCCTGGGAGAAG
TCACTGTGGACAAGGTTCTCGAGTTCCCATTGCTGAAAGGCCACCTAACCTCTGTGTTACAGGTG
AACAGCTCGGAGAACTGCAGTGCAGGAGTGCTTGTGGGCTGACTCCCTGGGGCTCTCAAAGA
CGATGTGCTACTGTTGGAGACACGAGGCCCCCCAGGACACAGATCCCTCTGTGCTTGGAAACCA
GTGGCTGTACTTCACTACCCAGCAAGGCTCCACGAGGCGAGTGCCTTGGAGAGTACTTACTA
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCTG
CCCCATGGACAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTTGCCG
CTGCGCTTCCCTCATCTTCCCTTCTCAAAAGGATCACGCGAAAGGGTGGCTGAGGCTTGTGAA
CAGGACCTCCGCTCGGGGCTCGCTCCAGGGCCGCGCGCTACTGCTACTACGCTCAGGTAGA
CTCGGGTTTCGAGCGCTGTTGGGCGCCCTGGCGTTCGGCCCTGTGCCAGCTGCCGCTGCGCGTGG
CCGTAGACCTGTGGAGCCGTCTGTAAGTGTGAGCGCGCAGGGGCCCGTGGCTTGGTTTCACGCGCAG
CGGCGCCAGACCTCGAGGAGGGCGCGGTGGTGGTCTTGCTCTTCTCTCCCGGCTGCGGTGGCGCT
GTGCGAGCGAGTGGCTACAGGATGGGGTGTCCGGGCCCGGGCGACGGCCCGCAGCGCTTCC
CGCTCTCGCTCAGTGTGCTGCTGCCGACTTCTTGAGGCGCGGCGCGCGCAGCTACGTGGG
GCTGCTTCGACAGGCTGCTCCACCCGACGCGGTACCCGCCCTTTTCCGACCGTGCCTGCTT
CACACTGCCCTCCCAACTGCCAGACTTCCTGGGGGCCCTGCAGCAGCTCGCGCCCCCGGTTCG
GGCGGCTCGAAGAGAGAGCGGAGCAAGTGTCCGGGCCCTTCAGCCAGCCCTGGATAGCTACTTC
CATCCCCCGGAGTCCCCGCGGGGACGCGGGGTGGGACCAAGGGCGGGACCTGGGGCGGGGA
CGGACTTAATAAAGGCAGACGCTGTTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSELRERLVGPQDATHCSPGLSCRLWSDILCLPGDIVPAPGPVLAPTHLQTELV
 LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVLSFQAYPTARCVLLEV
 QVPAALVQFGQSVGSSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSAADGNVHLVLNV
 EEQHFGLSLYWNQVQGGPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQ
 LWQAARLRLTLQSWLLDAPCSLPAEAALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFLLKGHPNLCVQ
 VNSSEKLQEQECLWADSLGPLKDDVLLLETRGPQDNRSLEPSGCTSLPSKASTRAARLGEYLLQDLQS
 GQCLQLWDDDLGALWACPMKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARG
 RAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAFWHAQRRQTLQEGGVVLLFSP
 GAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLHPDAVPALERTVPVFT
 LPSQLPDFLGALQQPRAPRSGRLQERAEQVSRAALQPALDSYFHPPGTFAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
 334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
 692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGAAGTGTGGGATCCCTGGCT
GCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCAACTTTGA
AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTGACGGGATCACCCGGAAGTCCTGC
AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
GTCAGTGCGGGAGGCGGTGAGCCACCAAGATGACTGACAGGTTGAGTCTCTGACGACACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGAGATGATTGTTTCATCCTA
CCCCACGCCAATCCGTGCAGGCGATGGCCACCGGCTAACCTGGAAGACATCTTCCATGACCTG
TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAGCAGAGAGA
ATATGAGTTCTTCGGCCTGACCCCTGACACAGAGTTTCTTGGCACCATCATGATTGCGTTCCCA
CCTGGGCCAAGGAGAGTGCCCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
TACTCCTTCTCCGGAGCCTTCTGTCTCCATGGGCTTCTCGTCGAGTACTCTGCTACCTGAG
CTACAGATATGTACCAAGCCGCTGCACCTCCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
TCCAGCCGCTGCGCTTCATCCAGGAGCAGTCTGATCCCTGTCTTTGACCTCAGCGGCCCCAGC
AGTCTGGCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
CCTCCAACGTGCCACCTCCCAAGATCCTCTCCCACTGTCTATGCCCAAAGCTGCCCTGAG
GTGGGCCCCCATCTATGCACCTCAGGTGACCCCGAAGCTCAATTCCCATTTACGCCCCACA
GGCCATCTCTAAGTCCAGCTTCTCCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
CCTATGGGGTATGATGGAAGGTTCTGGCAAAGACTCCCCCACTGGGACACTTTCTAGTCTAAA
CACCTTAGGCCTAAAGGTGAGCTTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT
TTCTCTGCAGGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAAGCAAAATCATTGCACCAAGC
CCCTGGGGATTGTCACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGGACA
CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC
CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
TGCTGGAGTCCCTTGTGTGTCCTCAAGGATGAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTG
GAGCAGCCACAGAACTGGATTCTCTTTTCAAGGCTGGCCCTGACTGTGCAAGTGGGAGTCCCTG
AGGGGAATGGGAAAGGCTTGGTGCTTCCCTCCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCA
ATCCCATGCCTGCCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC
AGAGGGAGTGGCATGCAGGGCCCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
AGGACTGCAGCGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCGTGTGCTGAGCCCTG
CAAGGCAGAAATGACAGTGAAGGAGGAAATGCAGGGAACCTCCCGAGGTCCAGAGCCCCACCTC
CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCTCTCCTTGCCCCATTCTGGCCAGTTTC
ACAATCTAGCTCGACAGAGCATGAGGCCCCCTGCCTCTTCTGTGCTATTGTTCAAAGGTGGGAAGAGA
GCCTGGAAAAGAACCAGGCCTGGAAAAGAACCAGAAGGAGGCTGGGCAGAACAGAACCACTGC
ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGCTCA
TTCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTGAGCTTCATTCTCTGATAGAACAAAGC
GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGACCCCT
ATCCTGAGAATGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
TGTAAGTGTGACAACTTTGCAAGCTCTGCCTTGGGTTGAGCCATCTGGGCTCAAATTCAGC
CTCACCCTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGGCCAGAACCTCGGTTTCCCT
ATCTGTAATGTGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
TCTTTAAAGTGCTTAATAGTGCCTGGTACATGGGCAGTGGCCAAATAACGGTAGCTATTTAAAAA
AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPD TVYSIEYKTYGERDW
 VAKKGCQRITRKSCNLT VETGNLT ELYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS
 KVRSIQMI VHTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMH LGGKQREYEFFGLTPDT
 EFLGTIMICVPTWAKESAPYMC RVKTL PDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
 PNSLNVRVLT FQPLRFIQEHVLIPVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
 YLGQPDISILQPSNVPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
 APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSLAM
 EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC
 SPSDQGPSPWGLLSLVCPKDEAKSPAPETSDLEQPTELDSLFRGLALT VQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCC'TACT'GGAAAAAAAAAAAAAAAAAAAAAGT'CACCCGGGCCCCGCGGTGGCCACAACATGG
 CTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG
 TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
 GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA
 AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAAC'TTTGGGCTGGA
 AGTGTTGAACACAGTTTGGATATTTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA
 AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT
 TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAAGTGGAGGACTCTGTACCTGAAGAG
 TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
 ACTTGACCCTGTGCCTGAGCCCAGGCATTCAGAGCTGATTCAGAGGATGGAGAAGGTGCTTTCT
 CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT
 CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCTGCACGATAA
 ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
 AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
 CATTACAGCAAAGGATTTTCGTTGGCATCAAATCTAAGTTTGT'TTTACAAAGATTGTTTTTTAGTA
 CTAAGCTGCCTTGGCAGTTTGCATTTTGTAGCCAAACAAAAATATATTATTTTCCCTTCTAAGTA
 AAAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFLVGLALWWVPGQSDLSHGRRFSDLKVCGDEECSSMLMYRGKALEDFTGPDRCRVN
 FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPAETDFVCFEGGRD
 DFNSYNVEELLGSLELEDSVPEESKKAEEVSQHREKSPPEESRGRELDVPPEPEAFRAEDGEDGA
 FSESTEGLOGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
 EKTDAYKVLKTEMSQRGSGQCVIHYSGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCAGAGGCCGGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCCTTAGGACATACACGGGACCCCCTAACCTC
AGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAACCTCAGCCCCGACATCCACGCGCGGCACAGG
CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTCGGGCAGCTGGGCTCGGGC
GGCGGGAGTAGGGCCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG
GGCAGAGGGCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGCATGAGCCGCTGG
TCTCGCTGCTGCTGGGCGCCGCTGCTCTGCGGCCACGGAGCCTTCTGCCGCCGCGTGGTCAGC
GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAC
GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCCTGGCTTGTGAGAGTGAGGGAGGAGTCCCTCTCA
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAACATCTGG
TGCTTGCCAGATCTTACCAGTGGTCTGATGGAAGCAATCCCAGTACCGAAACCTGGTACACAG
ATGAACCTTCTGCGGAAGTGAAGTGTGTTGTGATGTATACCAACCAACTGCCAATCCCTGGC
CTTGGGGGTCCCTACCTTTACCAGTGGAATGATGACAGGTGTAACATGAAGCACAATTATATTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCCGTAGAAAAGCCTTATCTTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTGTACTGAAGCAGGTATAATTCCTAATCTAATTTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTGTGCTTTTGGAACTGTTGTTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAACAACTAGTCCAAACAGTCTACACTGTGGATTTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTTCAGAAATTTTGTAAATCT
GGATCTGTATAAGGAATGGCATCAGAACAATAGCTTGAATGGCTTGAATCACAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTTGAGGCAATATTAAGTAATTTTATATGTCTATTATTTC
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGTAAAGGATGCACCCAA
ACTTCAAACCTTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
TGTGTGTTAGAAGCAATTCCTTTTATTTCTTTACCTTTTCATAAGTTGTTATCTAGTCAATGTAA
TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGTTGATAAAA
ATGAACCTGTTCTAATATTTATTTTATGGCATCTCATTTTCAATACATGCTCTTTGATTAAAG
AACTTATTACTGTTGTCAACTGAATTCACACACACACAATATAGTACCATAGAAAAAGTTTGT
TTTCTCGAAATAATTCATCTTTACGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA
AATAAGAAGCTATTTTATTAAGTGTGATATAAACCCTCCTCAAACATTTTACTTAGAGGCAAGGAT
TGTCTAATTTCAATTTGTGCAAGACATGTGCCTTATAATATTTTGTAGCTTAAATTAACAGATT
TTGTAATAATGTAACCTTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAA
TGACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTGA
GGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTCGGGGTTTGGG
ATTGACACTGGAGGCAGATAGTTGCAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG
ACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACACGAAATCGTGTGAAAATGGGTTGG
AACCCTCAGTGATCGCATATTTCATTGATGAGGGTTTGGCTTGAGATAGAAAATGGTGGCTCCTTT
CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCCTTGTCTCTCAAGAGAAAGTTGTAACCT
CTGGTCTTCATATGTCCCTGTGCTCCTTTTAACCAATAAAGAGTTCTTGTCTTGGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

[illegible]

Important features of the protein:

amino acids 1-21

amino acids 214-235

amino acids 86-89, 255-258

amino acids 266-269

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217